



Escuela de
Ingeniería y Arquitectura
Universidad Zaragoza



Modelo de caracterización del espacio (simetrías y dimensionalidad) en agentes artificiales a partir de sensores y efectores desconocidos

PROYECTO DE FIN DE CARRERA

Autor: Jorge Ochoa Villar

Director: Manuel González Bedia

Codirector: Francisco Serón Arbeloa

Ingeniería en Informática
Curso 2013-2014

Departamento de Informática e Ingeniería de Sistemas

Escuela de Ingeniería y Arquitectura

Universidad de Zaragoza

Junio de 2014

Modelo de caracterización del espacio (simetrías y dimensionalidad) en agentes artificiales a partir de sensores y efectores desconocidos

RESUMEN

En este proyecto nos hemos basado en ideas desarrolladas en modelos mínimos de percepción espacial generada a partir de la extracción de información en interacción sensomotora. Se revisa en el trabajo un modelo matemático que permite formalizar el modo en que un agente podría deducir las propiedades espaciales del mundo que habita de una manera sensomotora, y, además, la posibilidad de un algoritmo que permita diseñar organismos artificiales sin imponerles el mismo hábitat espacial que el de su diseñador.

Basándonos en estas ideas, se ha desarrollado un entorno experimental con el fin de testear las ideas que sustentan el modelo y su implementación matemática.

El marco experimental ha consistido en el diseño de un entorno virtual con un agente sensor-motor que se relaciona con el entorno a través de unas leyes psicofísicas que se han implementado basándose en descripciones básicas provenientes de organismos biológicos.

Además, se ha diseñado un entorno experimental físico con un agente (robot *Lego Mindstorms NXT*) que se relaciona con su entorno del que recibe estímulos visuales, probando la robustez del modelo en entornos reales.

Índice general

I	Memoria	IX
1.	Introducción	1
1.1.	Objetivo y alcance del proyecto	1
1.2.	Contexto en el que se realiza el proyecto	1
1.3.	Trabajo a realizar	2
1.4.	Herramientas utilizadas	2
1.5.	Estructura del documento	3
1.6.	Planificación	3
2.	Estado del arte	5
2.1.	Leyes físicas, psicofísicas y sensomotoras	5
2.2.	Enfoque constructivo de la percepción espacial: descubriendo el espacio desde las dependencias sensomotoras	6
2.3.	“Comprensión es comprensión”	7
2.4.	Modelos de cognición mínima	8
3.	Modelo formal	11
3.1.	Introducción	11
3.2.	Nociones fundamentales en procesos sensomotores de percepción espacial	11
3.2.1.	Condiciones iniciales	11
3.2.2.	Desarrollo de la capacidad perceptiva del organismo	12
3.2.3.	Hacia una comprensión de la estructura del espacio	13
3.3.	Presentación del modelo matemático	14
3.3.1.	Grupo de isotropía de la ley sensomotora	15
3.3.2.	Campo de vectores fundamentales sobre la variedad S	17
3.3.3.	Órdenes motoras generadoras en X^S	19
3.3.4.	Estructura de $\Gamma^d(\varphi)$ a partir de X^S	20
3.4.	Algoritmo	22
4.	Entorno experimental	25
4.1.	Diseño del entorno virtual y del agente sintético	25
4.1.1.	Sistema sensor del agente	26
4.1.2.	Sistema motor del agente	28
4.1.3.	Sistema nervioso del agente	29
4.2.	Experimento tipo 1: Entorno simulado	30
4.2.1.	Aplicación del algoritmo	30
4.2.2.	Resultados	32
4.3.	Diseño del entorno real y del agente físico	34
4.3.1.	Evaluación de la viabilidad experimental	35
4.4.	Experimento tipo 2: Entorno real	39
4.4.1.	Configuración experimental	39

4.4.2. Diseño experimental	40
4.4.3. Ejecución experimental y resultados	42
5. Conclusiones	45
5.1. Objetivos alcanzados	45
5.2. Trabajo futuro	45
II Anexos	47
A. Nociones matemáticas	49
A.1. Nociones de topología, geometría y geometría diferencial	49
A.2. Simetrías en transformaciones geométricas, noción de grupo y representación matricial	51
A.2.1. Estructura de grupo	52
A.2.2. Representación matricial	53
A.3. Nociones de Grupo de Lie y álgebra de Lie	54
A.4. Algebra de Lie asociada a un grupo de Lie	55
A.4.1. Método operativo entre álgebras y grupos de Lie	56
A.5. Estructura de un grupo de Lie capturada a través de su álgebra	57
A.5.1. Algebras de Lie asociativas y métricas sobre un álgebra	59
A.6. Ortonormalidad de una base en un álgebra de Lie	61
A.6.1. Propiedades de las métricas	61
A.6.2. Métrica en un álgebra de Lie	65
B. Algoritmo	67
B.1. Implementación del algoritmo [11]	67
B.1.1. Differential - Algoritmo de cálculo del espacio tangente	68
B.1.2. Svd_differential - Algoritmo de bootstrapping diferencial	69
B.2. Resumen de los detalles de implementación del entorno virtual	70
B.2.1. Estructuras matriciales	70
B.2.2. Resumen de las leyes psicofísicas asociadas a los dispositivos sensores	71
B.3. Detalle de las etapas del algoritmo	72
B.3.1. Etapa 1: Condiciones iniciales	72
B.3.2. Etapa 2: Obtención de una base de $T\varphi(M, E_0)$	72
B.3.3. Etapa 3: Obtención de una base de $T\varphi(M_0, E)$	73
B.3.4. Etapa 4: Obtención de una base de $T\varphi(M_0, E) \cap T\varphi(M, E_0)$	74
B.3.5. Etapa 5: Obtención de una base de vectores generadores de X^S	74

Índice de figuras

1.1. Diagrama de Gantt de las actividades realizadas.	3
2.1. Ejemplo de vehículo de Braitenberg	6
3.1. Organismo sensomotor.	12
3.2. Organismo interaccionando con su entorno.	12
3.3. Movimiento compensable. El organismo no percibe ningún cambio en sus sensores.	13
3.4. Ejemplo de espacio tangente y mapeo exponencial.	17
3.5. Extracción de espacios tangentes $T\varphi(M_0, E)$ y $T\varphi(M, E_0)$	19
3.6. Espacio tangente intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$	20
4.1. Agente simulado con tres tipos de dispositivos sensomotores.	25
4.2. Ejemplo esquemático del funcionamiento del dispositivo visual. La proyección de las fuentes del entorno se modelan siguiendo el modelo de cámara estenopeica.	26
4.3. Ejemplo esquemático del funcionamiento de la retina. El total de intensidad luminosa que recibe el fotorreceptor se modela mediante una gaussiana en función de la distancia entre el fotorreceptor y la proyección luminosa sobre la retina.	27
4.4. Ejemplo esquemático del funcionamiento del dispositivo auditivo. La antena favorece (a) las fuentes auditivas situadas frente a los dispositivos sensoriales y penaliza (b) los cambios de orientación entre fuente y antena.	28
4.5. Ejemplo esquemático del funcionamiento del dispositivo auditivo. El agente percibe el desplazamiento de su flagelo debido a las perturbaciones provocadas por la fuente.	28
4.6. Ejemplo esquemático del sistema nervioso sensor antes (a) y después (b) de la auto organización neuronal.	29
4.7. Ejemplo esquemático del sistema nervioso motor.	30
4.8. Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M, E_0)$	33
4.9. Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M_0, E)$	33
4.10. Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M_0, E) + T\varphi(M, E_0)$	34
4.11. Primer prototipo de agente sensomotor. En este ejemplo en particular, los sensores colocados en los extremos de los brazos eran unas cámaras web.	36
4.12. Ejemplo de prueba realizada con cámara web (a) y detalle del entorno (b).	37
4.13. Detalle del entorno.	38
4.14. Robot con Kinect.	39
4.15. Entorno final de simulación.	40
4.16. Detalles de la simulación de la proyección de los elementos del entorno sobre la retina.	41
4.17. Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M, E_0)$. Como puede observarse en la figura, el número de iteraciones realizadas en esta etapa del algoritmo es de 4 iteraciones, frente a las 8 del experimento virtual (ver Anexo B). Esta reducción se programó así porque se vio en los resultados que 4 iteraciones eran necesarias para caracterizar el espacio $T\varphi(M, E_0)$ correctamente.	43
4.18. Resultados de la simulación adicionales, correspondientes a las fases 3 y 4 del algoritmo.	44

A.1. Transformaciones topológicas.	49
A.2. Campo de vectores sobre la superficie de una esfera.	51
A.3. Ejemplo de función exponencial.	58
A.4. Espacios tangentes de $O(2)$ y $SO(2)$	58
A.5. Proyección de un cilindro en \mathbb{R}^3 sobre \mathbb{R}^2	65

Índice de tablas

4.1. Resumen de prestaciones y configuración del sistema motor.	35
4.2. Resumen de precisiones de sensores y motores.	36
4.3. Resumen de características de las cámaras probadas.	37
4.4. Resumen de resultados de las pruebas de precisión con el sensor de profundidad de la Kinect.	38
4.5. Resumen de los parámetros de configuración del entorno en los experimentos virtual y real.	40
4.6. Resumen de los parámetros sensores de los experimentos virtual y real.	42
4.7. Resumen de las configuraciones motoras y nerviosas de los experimentos virtual y real.	42
B.1. Algoritmos implementados por [11].	67
B.2. Descripción del algoritmo <i>differential</i>	68
B.3. Descripción del algoritmo <i>svd_differential</i>	69
B.4. Descripción de las estructuras matriciales utilizadas en la simulación.	70
B.5. Descripción de las leyes psicofísicas asociadas a los dispositivos sensores.	71
B.6. Descripción de la etapa 1 del algoritmo.	72
B.7. Descripción de la etapa 2 del algoritmo.	73
B.8. Descripción de la etapa 3 del algoritmo.	73
B.9. Descripción de la etapa 4 del algoritmo.	74
B.10. Descripción de la etapa 5 del algoritmo.	74

Parte I

Memoria

Capítulo 1

Introducción

1.1. Objetivo y alcance del proyecto

En este proyecto se pretende comprobar mediante experimentos en entornos virtuales y reales la validez de un modelo matemático de caracterización del espacio (simetrías y dimensionalidad) en agentes sensomotores.

Para ello, nos hemos basado en las ideas desarrolladas en [10, 11]. En estos trabajos, los autores proponen un modelo matemático que les permite formalizar el modo en que un agente podría deducir las propiedades espaciales del mundo que habita de una manera sensomotora, y, además, la posibilidad de un algoritmo que permita diseñar organismos artificiales sin imponerles el mismo hábitat espacial que el de su diseñador.

Hemos usado las ideas de estos autores y su implementación matemática, para construir un entorno experimental en el que poner a prueba el modelo.

El marco experimental ha consistido en el diseño de un entorno de simulación virtual con un agente sensomotor con tres tipos de sensores (visuales mediante “ojos”, auditivos mediante “antenas” y táctiles mediante “flagelos”) que se relaciona con el entorno a través de unas leyes psicofísicas (visual, auditiva y táctil) que se han implementado basándose en descripciones básicas expuestas en trabajos como [11, 7]. Suponiendo que el agente no tiene ninguna información a priori sobre el espacio físico en que actúa y que tampoco conoce la correspondencia entre su cuerpo y su sistema nervioso (es decir, es incapaz de realizar movimientos controlados), el experimento simula sus movimientos aleatorios y las respuestas sensoriales que recibe de su entorno como consecuencia a dichos movimientos. El objetivo de este experimento es comprobar si, a través del algoritmo matemático propuesto, el agente es capaz de hallar las propiedades de simetría y la dimensión del espacio en que actúa a través de estas relaciones entre sus movimientos y sus respuestas sensoriales.

Además, se ha diseñado un experimento en un entorno real físico para ver comprobar la validez del modelo propuesto en entornos reales con agentes reales. Para ello, se ha utilizado un robot *Lego Mindstorms NXT* que se relaciona con su entorno a través de una cámara *Kinect Xbox 360*, simulando la actuación de una ley psicofísica de naturaleza visual.

1.2. Contexto en el que se realiza el proyecto

En el ámbito de las ciencias cognitivas contemporáneas se está produciendo un desplazamiento creciente desde una perspectiva que considera que las capacidades cognitivas se procesan, ejecutan y desarrollan exclusivamente en el cerebro (dominada esta aproximación por la metáfora del ordenador), hacia una perspectiva en la que el enfoque recae sobre la interacción dinámica agente-entorno [5]. En particular, varios autores teorizan que la percepción que tenemos del espacio que nos rodea, en oposición a perspectivas que consideran que se basa en una representación interna del entorno,

depende de la forma en la que extraemos información implícita en los bucles de retroalimentación sensoriomotora que establecemos de manera automática [8].

Este enfoque se conoce como la teoría sensoriomotora de la percepción [9] e intenta desarrollar y fundamentar una alternativa a la teoría computacional de la percepción [14], mediante modelos formales y computacionales. Aunque han existido avances teórico-conceptuales en los últimos años en la teoría sensoriomotora de la percepción, estos no han venido acompañados de técnicas de modelización y experimentación capaces de cuantificar y testear la viabilidad de las hipótesis.

En la Inteligencia Artificial clásica, se define el comportamiento de los agentes artificiales siguiendo un enfoque motor, esto es, se programa un algoritmo que controla cómo se comporta el agente en función de los estímulos que recibe. Este algoritmo que se le proporciona al agente suele estar modelado de acuerdo a leyes físicas y psicofísicas del espacio en que actúa el agente (y que conoce el programador), lo cual requiere suponer que el agente posee “a priori” un conocimiento sobre las leyes de su entorno que rigen su comportamiento.

Sin embargo, en el enfoque sensoriomotor, el propósito del agente es el de extraer, precisamente, esas leyes, esos algoritmos que relacionan sus acciones con sus estímulos sensoriales. Y, además, lo hace observando cómo cambian sus sensores como consecuencia de sus acciones, al revés de como lo hace la inteligencia artificial clásica, donde son sus acciones las que son consecuencia de sus estímulos sensoriales. De esta manera, en el enfoque sensoriomotor, no es necesario un conocimiento “a priori” sobre las leyes físicas y psicofísicas del espacio en que actúa, pues es el propio agente el que construye, a través de sus relaciones con el entorno, sus propias “leyes sensoriomotoras” con las que interactúa. Aunque para que el agente sensoriomotor sea capaz de extraer dichas leyes, es necesario que disponga de algún mecanismo (o algoritmo) que le permita extraer dichas relaciones y construir sus propias leyes. El funcionamiento de dicho algoritmo se basará en relaciones sensoriomotoras, y no en un conocimiento “a priori” de las leyes físicas y psicofísicas del espacio en que actúa.

La teoría sensoriomotora de la percepción busca, por tanto, encontrar nuevos modelos (tanto teóricos como computacionales) que permitan a un agente descubrir sus propias relaciones sensoriomotoras a través de sus relaciones acción-percepción con su entorno, pero sin suponer ningún mecanismo a priori basado en el conocimiento de las leyes físicas y psicofísicas implicadas en dichas relaciones.

1.3. Trabajo a realizar

Para alcanzar el objetivo de este proyecto, que es el de poner a prueba tanto las ideas como el modelo matemático propuesto en [10, 11] mediante la construcción de unos experimentos en entornos de prueba tanto virtual como real, se realizarán las siguientes tareas:

1. Documentarse acerca de la teoría sensoriomotora de la percepción y adquirir una base de fundamentos matemáticos (geometría diferencial, grupos de Lie) necesarios para entender tanto el modelo conceptual como el algoritmo matemático propuesto.
2. Analizar la implementación matemática de [11] para poder integrarla en modelos de simulación tanto virtual como real.
3. Diseñar un experimento en un entorno simulado para evaluar el modelo matemático propuesto.
4. Diseñar un experimento en un entorno real con robots que permita comprobar la aplicabilidad del modelo en entornos físicos.
5. Evaluar y proponer posibles vías de trabajo futuro a partir del modelo propuesto.

1.4. Herramientas utilizadas

Las herramientas utilizadas a lo largo del desarrollo del proyecto han sido las siguientes:

- Entorno de programación para el diseño y ejecución de los experimentos:
 - Entorno MATLAB
- Construcción del agente sensomotor (robot y sensores).
 - Lego Mindstorms NXT
 - PlayStation Eye PS3
 - Hercules Classic Silver
 - Kinect Xbox 360
- Control del robot.
 - RWTH - Mindstorms NXT Toolbox for MATLAB

1.5. Estructura del documento

La estructura de esta memoria está dividida en cuatro capítulos, incluyendo este capítulo introductorio. En el capítulo 2 se definen los conceptos acerca del enfoque sensomotor sobre los que se basa este proyecto así como otras ideas acerca de la percepción. En el capítulo 3 se describe el modelo tanto conceptual como matemático sobre el que se basa el trabajo realizado y el algoritmo que lo implementa. En el capítulo 4 se describen los experimentos realizados y se analizan los resultados obtenidos. Por último, en el capítulo 5 se recogen las conclusiones extraídas de los experimentos realizados y se discuten los posibles trabajos futuros.

1.6. Planificación

Durante los 13 meses de duración del proyecto, se han realizado las tareas que se muestran en el siguiente diagrama de en el diagrama de Gantt contenido en la figura 1.1.

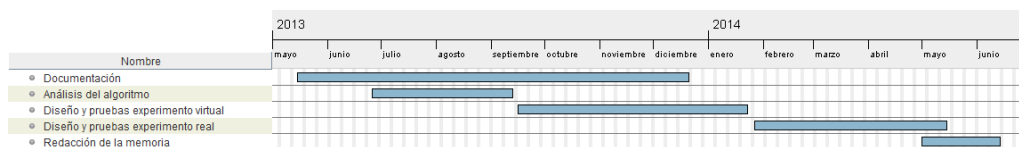


Figura 1.1: Diagrama de Gantt de las actividades realizadas.

Capítulo 2

Estado del arte

2.1. Leyes físicas, psicofísicas y sensomotoras

En las ciencias tradicionales, el fin de la investigación es la obtención de leyes naturales dirigidas a dar explicaciones causales: estas leyes se conocen como leyes físicas y se considera que capturan la estructura real de los fenómenos en la naturaleza. Son leyes diferentes a las denominadas leyes psicofísicas, que asumen un sujeto en interacción con un mundo, y que se encargan de representar cómo en el sujeto impactan los fenómenos naturales (a qué espectro luminoso son sensibles sus órganos de la visión, qué sensibilidad acústica tienen sus dispositivos auditivos, etc.).

En esta sección nos interesa introducir un tipo de ley diferente a las anteriores y que denominaremos *ley sensomotora*. Estar gobernado por una ley sensomotora es asumir que el sistema no es un ente que recibe información de manera pasiva. Al contrario, implica una relación entre la acción del sistema y su percepción que, a diferencia de los otros tipo de leyes, funciona de manera activa, es decir, el sistema extrae una serie de regularidades entre los movimientos que despliega y los estímulos que recibe. Una ley física (o psicofísica) es una ley fija, proporcionada desde fuera, sin posibilidad de intervención en la configuración de la misma por parte de un agente. Una ley sensomotora que regule el comportamiento de un agente, por contra, es una ley que construye el agente en la propia interacción que establece con su entorno, de manera que la acción es necesaria para la percepción y, a la vez, es la percepción la que condiciona la acción del sujeto, estableciendo bucles de causalidad circular.

Para poder diferenciar las nociones introducidas, consideremos un simple dispositivo fototáctico, por ejemplo, uno de los vehículos de Braitenberg [3]. El vehículo está equipado con dos sensores situados uno al lado del otro enfrente del vehículo y con dos ruedas situadas una al lado de la otra en la parte trasera. El sensor izquierdo está conectado al mecanismo que mueve la rueda derecha y el sensor derecho está conectado al mecanismo que mueve la rueda izquierda. Como resultado de estas conexiones, el vehículo se orientará por él mismo hacia las fuentes lumínicas y se moverá hacia ellas (ver Fig. 2.1). Un mecanismo tan simple puede detectar y perseguir fuentes luminosas.

Una *ley física*, presente en el ejemplo y relacionada con los estímulos, es la ley de propagación de la luz: la luz se propaga en un medio ideal en todas sus direcciones por igual en forma de onda (propiedad de isotropía) y su intensidad disminuye en una relación inversamente proporcional al cuadrado de la distancia al centro donde se origina.

En el ejemplo, también se puede hablar de *leyes físicas motoras*. Un ejemplo de ellas serían el conjunto de leyes mecánicas que explican el desplazamiento de las ruedas por el suelo a partir de la fuerza ejercida por el motor sobre la rueda, la fuerza de rozamiento con el suelo, etc.

Podemos distinguirlas de las correspondientes leyes psicofísicas sensoras y motoras asociadas. Por ejemplo, una *ley psicofísica sensora* caracterizaría la relación entre la magnitud de la intensidad luminosa que le llega procedente de la luz al vehículo, y la intensidad con la que ésta es percibida por el sensor del vehículo. En este caso, por ejemplo, los sensores podrían registrar la intensidad

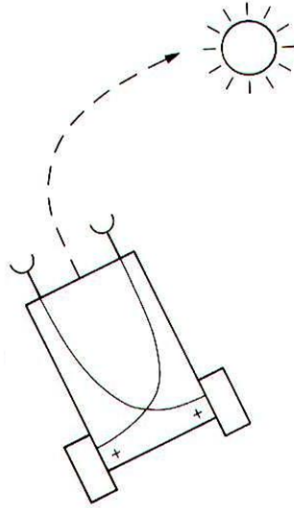


Figura 2.1: Ejemplo de vehículo de Braitenberg

luminosa siguiendo una función logarítmica con unos umbrales superiores e inferiores a partir de los cuales el sensor deja de detectar la luz.

La correspondiente *ley psicofísica* motora del ejemplo establecería la relación entre la señal que le llega al motor procedente de los sensores y la acción que éste ejerce sobre la rueda. En este caso, la ley dependería de la precisión del motor a la hora de convertir la señal recibida en potencia de giro, del umbral de señal en el que el motor es capaz de funcionar, etc.

Por último, en nuestro ejemplo, la *ley sensomotora* es la ley que relaciona la acción del vehículo con los estímulos que recibe, es decir, que combina señales sensoras y motores. Si, como ocurre con la dinámica de los vehículos de Braitenberg, hay una fuente luminosa a la izquierda del vehículo, esto hará que el vehículo gire en la dirección de la fuente luminosa. Una vez que el vehículo ha girado, la fuente luminosa ya no estará a su izquierda, sino a su derecha. Esto causará que el vehículo gire hacia la derecha hasta que, de nuevo, la luz vuelva a quedar a la izquierda. Esta serie de relaciones estímulo-acción irán moviendo al vehículo hacia la fuente luminosa. De esta manera, la estimulación sensorial que el sistema recibe es dependiente de las acciones que realiza, y, a su vez, las acciones que realiza también son debidas a los estímulos que recibe. Podemos decir que un vehículo así construido despliega una dinámica gobernada por una ley sensomotora, es decir, un conjunto de reglas de interdependencia entre estímulos y movimiento que le permiten, en este caso, perseguir fuentes luminosas.

2.2. Enfoque constructivo de la percepción espacial: descubriendo el espacio desde las dependencias sensomotoras

El enfoque constructivo del espacio [12], es decir, que la percepción espacial se construye en los primeros meses de vida, es una de las aportaciones más relevantes de los estudios sobre desarrollo cognitivo (de igual manera, la teoría piagetiana considera que el pensamiento formal, las operaciones lógicas, las reglas de conservación..., todas son capacidades que surgen en el pensamiento infantil a partir de la interacción sensomotora). La realidad que percibe el sujeto es fruto de su propia construcción intelectual.

Este enfoque sensomotor es contrario a la percepción realista del mundo (que adoptan las ciencias naturales como la física) que entiende que la relación entre el sujeto y su entorno, vía la

percepción, se produce en términos de mecanismos de representación. Para el realismo, la realidad externa existe independientemente de nuestro proceso de percepción, y éste se reduce a un proceso de representación en forma de estructuras mentales simbólicas. La forma de la realidad que percibimos no se ve afectada por el hecho de cómo la conocemos: la representación mental es una copia de la realidad.

Esto nos llevará a diferenciar conceptualmente dos tipos de mundo, en función del enfoque que adoptemos: (i) en el enfoque realista, el mundo en que habita un agente es el conjunto de fenómenos físicos que vienen caracterizados por leyes naturales. A este mundo lo denominaremos “entorno” del agente; (ii) en el enfoque constructivista, el mundo del agente está construido por él a través de las interacciones sensomotoras que establece con el medio. A éste, lo denominaremos el “espacio sensomotor” que percibe el agente.

El concepto de “espacio sensomotor” fue planteado por primera vez por el biólogo Jakob von Uexküll [16], introduciendo la noción de Umwelt (en alemán, “medio ambiente”). El Umwelt es el conjunto de todas las relaciones significativas del ser vivo con su entorno. Según von Uexküll, cada ser vivo tendrá un Umwelt, un mundo propio diferente que creará e incluso modificará según interaccione con su entorno. El concepto de Umwelt escapa de la idea de un mundo común para todos los seres vivos regulado por leyes físicas para llegar a la noción de un mundo propio y único que crea cada ser vivo según interacciona con su entorno. Es interesante, por tanto, destacar que la noción de Umwelt puede relacionarse con el enfoque sensomotor, pues es mediante la acción y la percepción como se crea y se modifica el Umwelt propio de cada ser vivo.

El enfoque sensomotor sobre la percepción [9] que se ha desarrollado en los últimos años, inspirada en las teorías de Piaget, considera que, en realidad, “percibir” mediante leyes sensomotoras nos aleja de la idea de que esta capacidad se base en el hecho de poseer mecanismos cerebrales preestablecidos específicos. En vez de eso, la teoría sensomotora propone que el hecho de “percibir” se basa en la interacción del cuerpo con su entorno, extrayendo una serie de regularidades, una estructura relacional, de esas interacciones. Para poder realizar dichas interacciones es obviamente necesario disponer de alguna estructura interna que haga de guía para poder realizar dichas interacciones pero sin depender del tipo de estímulo (el mismo mecanismo sirve para percibir estímulos visuales o táctiles, por ejemplo). En este caso, aunque existen mecanismos neuralmente codificados, sin embargo, su dominio no se encuentra sólo en el cerebro, sino, como se ha recalcado, en la actividad perceptual misma.

Esta nueva perspectiva propone que lo que determina la modalidad sensorial experimentada de una estimulación concreta viene dada por la ley sensomotora que gobierna dicha estimulación y no por el estímulo concreto. Según esto, lo que determina la naturaleza particular de la estimulación (visual, táctil, auditiva, olfativa, etc.) no es nada directamente relacionado con las vías sensoriales o las zonas cerebrales que llevan el influjo nervioso. Habrá leyes sensomotoras para cada tipo de acción perceptiva (por ejemplo, la experiencia táctil estará determinada por los cambios en los receptores sensoriales que ocurren cuando movemos el cuerpo con respecto a una superficie, mientras que la percepción auditiva estará determinada por leyes como el hecho de que cuando te aproximas a una fuente sonora, la intensidad del estímulo sensorial aumenta). Este resultado proporciona un marco unificador entre las diferentes modalidades sensoriales (como la sinestesia táctil) y la posibilidad de sustitución sensorial. Por ejemplo, si a través del tacto reproducimos condiciones similares a las que son usualmente obtenidas por los caminos visuales, éstas serán percibidas como si poseyeran una naturaleza visual (debido a que son gestionadas por la misma ley sensomotora). De hecho, hay evidencia de que esto es posible [1].

2.3. “Comprensión es compresión”

La forma de disponer de una estructura que caracteriza el Umwelt de un agente a partir de sus relaciones sensomotoras, implica necesariamente algún mecanismo de procesamiento de información que, de todo el mar de estímulos que recibe el sistema, sepa extraer y conservar sólo aquellas realmente significativas en su interacción con el mundo.

Esto es, el sistema “comprenderá” el mundo que habita y sus propiedades, en el momento en que sea capaz de “comprimir” la información recibida, siendo capaz de entenderla y categorizarla mediante un simple conjunto de reglas que la expliquen.

De acuerdo con [17], en la práctica de la Ciencia, podemos diferenciar tres formas diferentes de comprender los hechos científicos: (i) modelos basados en la causalidad (entender los fenómenos en términos de la observación de relaciones causa-efecto, como por ejemplo, la física), (ii) modelos basados en la estructura (entender a través de la síntesis y el análisis de las diferentes partes en las que puede dividirse el estudio de un fenómeno, como por ejemplo, la biología) y (iii) modelos basados en la comprensión (entender un fenómeno mediante la reducción de su complejidad, como en Ciencias de la Computación).

El enfoque basado en la comprensión asume una visión algorítmica de los fenómenos que nos permite conectar ideas provenientes de disciplinas como la Educación o la Psicología, y tratarlas en forma de modelos computacionales que, eventualmente, permiten aplicarlas al diseño de organismos artificiales.

En [4], el autor propone que la comprensión es un tipo de compresión de datos. Mantiene que “entendimiento es compresión y que una teoría científica nos permite comprender el mundo porque unifica muchos fenómenos mostrando que estos reflejan un mecanismo interno común”. Cuanto más simple es una teoría sobre un fenómeno, más profundo es el entendimiento sobre él. La comprensión es el límite de una conceptualización, y comprender algo es haberlo conceptualizado a una medida determinada.

Tal como se ejemplifica en [4], comprender algo es poseer un algoritmo que permita comprimir la información recibida y convertirla en descripciones algorítmicas más concisas. Por ejemplo, entendemos el número 0,33333... por pensar en él como un tercio. La primera forma de representar el número, requiere una cantidad infinita de memoria, mientras que la segunda forma puede producir todos los datos de la primera representación utilizando mucha menos información.

2.4. Modelos de cognición mínima

Dentro de las Ciencias Cognitivas, la cognición tiende a ser interpretada desde una perspectiva antropocéntrica, es decir, proyectando esquemas que provienen de nuestros prejuicios sobre cómo funcionan las capacidades humanas. Para evitar estos prejuicios a la hora de elaborar teorías surgió hace una década la línea de investigación conocida como “cognición mínima” [2, 15]. La cognición mínima busca modelos que, con mínimos recursos, sean capaces de materializar capacidades cognitivas primitivas o los orígenes de capacidades cognitivas más elaboradas. En la última década, la relevancia en la comprensión de fenómenos cognitivos a partir de modelos mínimos ha ido en aumento.

Concentrarse en una versión simplificada, mínima, de un fenómeno de estudio, sirve al investigador como herramienta para pensar sobre el fenómeno. Por ejemplo, algunos modelos deliberadamente simples pueden confirmar que existen alternativas explicativas a fenómenos que se suponía que funcionaban de una determinada manera. En palabras de algunos autores [2, 13], los modelos sirven “como gimnasia mental”, es decir, son entidades equivalentes a los experimentos mentales clásicos, son artefactos que nos ayudan a pensar.

En el área del estudio ingenieril de la inteligencia, podemos encontrar dos motivaciones: (i) La primera es puramente pragmática. Trata de encontrar algoritmos útiles en campos industriales, tecnológicos, etc., que tienen una amplia gama de beneficios. (ii) La segunda motivación es científica, esto es, se usan computadores, modelos de simulación, bots, etc. como plataformas experimentales para la investigación de cuestiones acerca de la inteligencia. Esta es la conocida como “metodología sintética” que se presenta como alternativa a la tradicional, y que defiende que el estudio ingenieril sobre la mente debería funcionar más como un campo científico que permita examinar hipótesis particulares, y en el que haya una investigación teórica para averiguar por qué ciertos modelos son mejores o no en términos de las asunciones que suponen.

Los modelos mínimos consisten en agentes sintéticos que habitan entornos virtuales y con ellos se pretende capturar aspectos concretos de una tarea cognitiva. Su condición de “mínimos” permite que puedan ser estudiados de manera completa. Estos modelos no son estructuras que organizan datos experimentales. Su valor proviene del valor científico que tiene el estudio de los patrones observados y la relación de tales patrones con hipótesis teóricas. Sirven como herramientas para cuestionar preconcepciones sobre como un cierto comportamiento es generado. Son modelos con una complejidad manejable que permiten un análisis completo y muchas veces analítico. El “análisis con modelos mínimos” en Ciencia Cognitiva aporta, en ocasiones, reflexiones que ponen en entredicho la validez de presupuestos metodológicos. Es decir, el análisis de modelos permite revelar las ideas ocultas al abordar problemas con una metodología determinada, impidiendo que existan explicaciones en otros términos.

Tal como hemos visto, podemos entender cómo un agente, gracias a sus interacciones sensomotoras, es capaz de definir su propio Umwelt, construyéndolo a través de las relaciones que establece con su entorno de una manera activa. Nuestro interés se centra en determinar cómo es posible que un agente pueda obtener información sobre la estructura de su espacio sensomotor planteándolo en términos algorítmicos, es decir, buscando un modo de formalizar estas ideas en una herramienta no sólo que permita comprender cómo el agente logra esto, sino también para disponer de mecanismos algorítmicos que permitan diseñar sistemas artificiales basados en estas ideas.

Capítulo 3

Modelo formal

3.1. Introducción

En este capítulo se explicarán las ideas que sustentan el proyecto, tanto a nivel teórico-formal como desde el punto de vista de su aplicación.

En una primera parte, se mostrarán las nociones fundamentales que explican cómo un organismo estructura el mundo en el que habita a partir de las relaciones operativas que establece con él. Se defiende aquí, por tanto, una visión constructiva de la capacidad de percepción en los organismos vivos (con el fin de que sirva para fundamentar otros modos de diseño de capacidades perceptivas en artefactos).

En una segunda parte, tras haber introducido los conceptos base, se pasarán a expresar matemáticamente estas nociones con el fin de poder disponer del problema en términos formales. A partir de la representación del problema obtenida, se pasará a explicar cómo, mediante técnicas que provienen del campo de: (i) la geometría diferencial, (ii) grupos de transformaciones invariantes y (iii) propiedades de álgebras de Lie, es posible diseñar un algoritmo que explote información sensomotora desplegada en el espacio de interacción entre un agente y su entorno, y la utilice para caracterizar las propiedades de invarianza y simetría del espacio en el que el sistema está embebido.

Las ideas teóricas de este capítulo y su implementación algorítmica están extraídas de los trabajos [10, 11].

3.2. Nociones fundamentales en procesos sensomotores de percepción espacial

Para explicar las nociones básicas, utilizaremos un agente sensomotor mínimo que nos permitirá introducir, de un modo ilustrativo, los conceptos necesarios sobre las que trabajaremos.

3.2.1. Condiciones iniciales

Supondremos un entorno compuesto por una serie de luces en el que habita un organismo simple consistente en un brazo articulado fijado al suelo con varias ramificaciones por cada articulación y con sensores sensibles a la luz en cada uno de sus extremos (véase Fig. 3.1). Se supone que posee un protocerebro mediante el que controla el movimiento de su cuerpo y procesa las señales sensoriales que recibe, a través de un sistema nervioso primitivo conectado a él. Suponemos que, inicialmente:

- El cerebro no tiene ninguna información a priori sobre el espacio físico en que habita.
- El cerebro no conoce la correspondencia entre su cuerpo y su sistema nervioso.

Como consecuencia de esto, el cerebro es incapaz de realizar movimientos controlados de su cuerpo, de tal manera, que lo único que puede hacer el organismo inicialmente es realizar movimientos aleatorios, recibiendo la correspondiente respuesta a dichos movimientos a través de sus sensores.

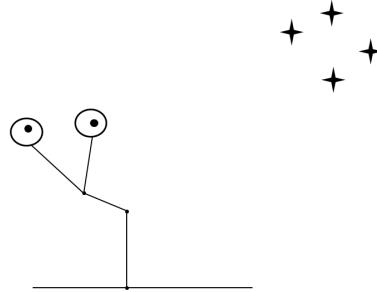


Figura 3.1: Organismo sensomotor.

3.2.2. Desarrollo de la capacidad perceptiva del organismo

Teniendo en cuenta estos movimientos aleatorios, y a partir de las diferentes respuestas sensoriales que recibe como consecuencia, el sistema es capaz de discriminar entre: (i) respuestas que presentan una relación sistemática con los movimientos que realiza, y (ii) respuestas que no parecen tener ninguna relación con los movimientos realizados. Esto le permite al organismo constatar que su universo tiene dos partes: (i) una que está directamente vinculada a sus movimientos (que llamaremos *cuerpo*) y (ii) otra que no presenta una relación aparente con sus movimientos (que denominaremos *entorno*). Una vez constatada esta diferencia, el organismo tiene la posibilidad de caracterizar ambas partes a través de la regulación de sus movimientos. Así, los cambios que se produzcan en los sensores cuando el cuerpo está estacionario serán los que provienen del entorno (ver Fig. 3.2b). Y a la inversa, si el cuerpo se mueve, el efecto en los inputs será debido al propio cuerpo¹ (ver Fig. 3.2a).

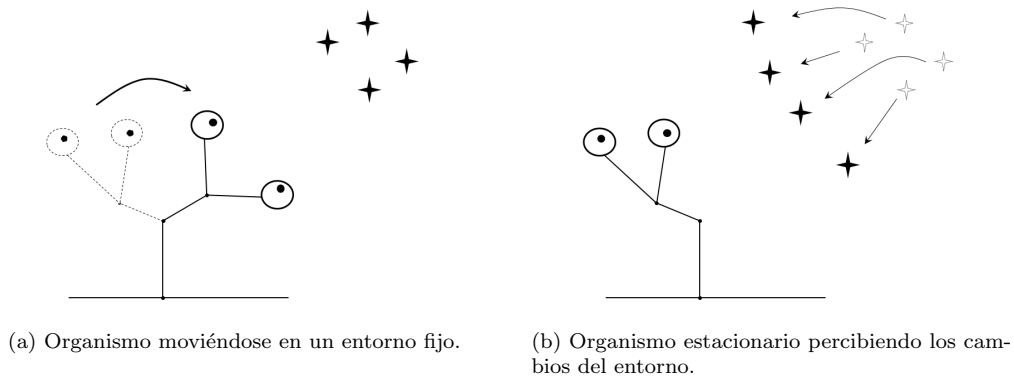


Figura 3.2: Organismo interactuando con su entorno.

¹En este caso podrá haber también una incontrolable contribución del entorno. Pero supondremos que los movimientos del organismo son más rápidos que los que genera la dinámica propia del entorno.

Supongamos que el organismo estructura estas correspondencias sensomotoras mediante algún mecanismo cerebral con el fin de hacer uso futuro y funcional de ellas (por ejemplo, mover su cuerpo controladamente en su entorno). Es por ello, que supondremos la existencia de un mecanismo de *eficiencia cerebral* que caracterizará estas nociones (cuerpo y entorno) en términos de una representación que utilice el menor número de parámetros posibles.

Por ejemplo, el organismo anterior es capaz de percibir la posición de una luz del entorno en términos del número total de sus sensores (por ejemplo 40) y caracterizarla de este modo. Sin embargo, sabemos que el menor número de parámetros necesario para representar un punto en el espacio es tres (sus coordenadas espaciales). Si existe un mecanismo cerebral en el organismo capaz de estructurar la información eficientemente (esto es, en un espacio de parámetros de dimensión menor que 40), éste será *máximamente eficiente* si encuentra una representación sólo en término de tres parámetros. Asumiendo este mecanismo cerebral, (1) al conjunto de correspondencias sensomotoras vinculadas al cuerpo del organismo una vez han sido estructuradas en el espacio de dimensionalidad mínima posible, lo denominaremos *representación del estado del cuerpo*; (2) análogamente, para el entorno percibido, tendremos una *representación del estado del entorno*.

Hay que recalcar que el organismo no tiene conocimiento “conceptual” a priori ni sobre sí mismo ni sobre el mundo en que habita. Estas nociones mencionadas (cuerpo “percibido” y entorno “percibido”) las construye el organismo a partir de las regularidades que detecta en sus procesos de interacción sensomotora con el mundo físico. Nótese que cuerpo y entorno “percibidos” no tienen nada que ver con las nociones de cuerpo y entorno físicos (ver sección 2.2).

Existe una diferencia conceptual entre ambas representaciones mencionadas: (i) la representación del estado del entorno es más básica, ya que no involucra el movimiento del organismo para obtener dicha representación; (ii) la representación del estado del cuerpo, sin embargo, es mucho más interesante, porque permite al agente conocer su propio cuerpo pero a través de los efectos que el entorno le proporciona cuando se mueve. Es decir, *el cuerpo “percibido” es una vía para descubrir las relaciones entre el cuerpo y el entorno*.

3.2.3. Hacia una comprensión de la estructura del espacio

Hay ciertos cambios en el cuerpo que pueden ser compensados perceptualmente por movimientos del entorno o del propio cuerpo. Por ejemplo, supongamos el organismo situado frente a una luz. Si el organismo se mueve uniformemente hacia la derecha, y la luz se mueve también uniformemente hacia la derecha y a la misma velocidad, ambos movimientos se compensarán perceptivamente. En esta situación, el organismo no habrá obtenido ninguna información relevante tras su interacción debido a que, a lo largo de este movimiento, la percepción de la luz en sus sensores se ha mantenido constante. A estas combinaciones de movimientos organismo-entorno que se compensan mutuamente se les denominan *movimientos compensables* (ver Fig. 3.3)

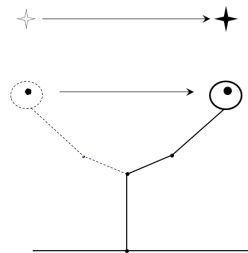


Figura 3.3: Movimiento compensable. El organismo no percibe ningún cambio en sus sensores.

La existencia de estos movimientos compensables implica una relación estrecha entre cuerpo y entorno. A partir de esta relación, podemos hablar de una “estructura común cuerpo-entorno” que

llamaremos *espacio sensomotor*².

Es importante darse cuenta de dos aspectos fundamentales sobre esta nueva estructura:

1. El organismo tiene acceso a ella desde un *enfoque sensomotor*, pues no existe la noción de compensabilidad para un observador pasivo. Ha sido mediante “acciones corporales” como el organismo descubre el acceso a un entorno distinto del cuerpo y de una estructura común que conecta a ambos³.
2. El organismo tiene acceso a ella en *términos de los movimientos (transformaciones) que la mantienen invariante*. Esto es, si el organismo se mueve haciendo observaciones sobre el efecto de sus interacciones, son aquellas que “no nota”, es decir, que no modifican sus sentidos, las que definen esta nueva estructura. Pensemos en el siguiente ejemplo: si el agente estuviese en el centro de una bóveda esférica y rotase, no detectaría ningún cambio en la distancia a la pared. Precisamente, porque las rotaciones son las transformaciones que dejan invariante a un espacio esférico. Nótese que el mundo (en este caso la bóveda esférica) queda expresado exactamente por el conjunto de transformaciones que lo dejan invariante (rotaciones con cualquier ángulo).

Nuestro objetivo es conocer la estructura del espacio sensomotor en términos matemáticos. Y, como acabamos de ver, el aparato matemático para su formalización debe proporcionar herramientas adecuadas que permitan expresar dicha estructura en términos de los movimientos que la mantienen invariante.

Con este enfoque se pretende dar una alternativa a un hecho asumido en el estudio de la geometría y que plantea una inconsistencia: cuando nos proponemos analizar la geometría del espacio en el que vive un agente (y definir una métrica) utilizamos implícitamente una idea a priori acerca de la existencia del espacio y asumimos que nuestro cerebro mapea los estímulos sensoriales en la idea arquetípica y abstracta de espacio que ya posee. En este trabajo el enfoque es diferente: cuando usamos términos como “representaciones del estado del cuerpo” describimos cómo nosotros (o nuestro cerebro) concibe el mundo pero sin ninguna relación con un modelo a priori (noción de Umwelt, ver sección 2.2).

3.3. Presentación del modelo matemático

En la Inteligencia Artificial clásica, definimos el comportamiento de los agentes siguiendo un “enfoque motor”, esto es, programamos un algoritmo F que implemente la relación

$$M = F(S, E)$$

es decir, programamos cómo se comportará el agente (comportamiento motor M) en función de los estímulos que reciba (S) y del estado del entorno en el que se encuentre (E).

En este trabajo, seguimos un “enfoque sensomotor”, en el que el propósito del sistema es ser capaz de deducir la función φ que articula la relación inversa

$$S = \varphi(M, E)$$

para ello el agente se moverá (M) en un entorno (E) y registrará cómo cambian los estímulos que recibe (S). A esta función φ la denominaremos *ley sensomotora* del agente.

Tal como hemos visto (ver sección 3.2.3), nuestro interés se centra en obtener la estructura del espacio sensomotor del agente en términos de las transformaciones que lo dejan invariante. Matemáticamente, esto nos lleva a modelar las estructuras del problema como variedades diferenciales (ver Anexo A). En geometría diferencial, una variedad S queda caracterizada a través de las propiedades de la variedad que se mantienen invariantes, bajo transformaciones de S .

²Esta estructura no es la del entorno físico real sino la estructura del espacio físico que percibe el agente (lo que hemos llamado Umwelt en la sección 2.2) [16].

³Experimentos como [6] introducen la importancia de la acción en la percepción.

Si en el enfoque clásico, “percibir” es algo que se presupone en los agentes (y que consiste en un proceso de recepción de información de manera pasiva y de representación interna), el enfoque sensomotor se pregunta por el hecho mismo de “percibir”. En concreto, cómo se crea la capacidad y las estructuras perceptivas al ser activadas ciertas terminaciones nerviosas por estímulos externos.

Pero el enfoque sensomotor, además, introduce un marco de estudio diferente: asumir que los organismos no perciben señales sensoriales puras sino “invariantes” en sus interacciones sensomotoras - esto es, “estructuras” ocultas en el mar de datos recibidos formados por parejas de órdenes motoras y retornos sensoriales – traslada el interés por conocer qué algoritmo seguiría el cerebro del agente, al interés por conocer cómo el cerebro es capaz de detectar estas estructuras invariantes ocultas en su ley sensomotora.

“Percibir”, por tanto, es capturar las propiedades intrínsecas de esa ley sensomotora.

Supongamos un agente con cuerpo P , en un entorno E . Su conjunto de movimientos motores los representamos por M y las señales sensoriales que recibe por S . Suponemos que la posición del cuerpo P está controlada por M a través de una función φ_a biyectiva,

$$P = \varphi_a(M)$$

que establece una relación directa $P \leftrightarrow M$ (denominamos a este tipo de sistemas holonómicos). Y suponemos que las señales sensoriales que recibe S dependen de la posición del cuerpo P y el estado del entorno, siguiendo una función φ_b

$$S = \varphi_b(P, E)$$

Ambas funciones pueden combinarse definiendo la ley sensomotora φ antes mencionada

$$S = \varphi(M, E)$$

Tal como hemos indicado antes, asumiremos que la estructura matemática de los elementos de la teoría M , S y E serán formalizados como variedades. Si suponemos que los cambios motores, los de entorno y los sensoriales son pequeños, y que las relaciones entre ellos son suaves, podemos asumir además que son variedades diferenciales. Consideraremos también que la función φ cumple las propiedades de derivabilidad que garantizan poder aplicar técnicas matemáticas de geometría diferencial.

3.3.1. Grupo de isotropía de la ley sensomotora

La misma ley sensomotora podría tomar diferente forma en dos agentes si sus sensores y motores son diferentes aunque su relación sensomotora con el mundo sea idéntica. Por ejemplo, dos agentes con sensores de diferente sensibilidad podrían presentar leyes diferentes

$$S = \varphi_1(M, E), S = \varphi_2(M, E)$$

pero ambas se corresponderían con la misma *ley sensomotora implícita*.

Nuestro objetivo no será conocer la ley sensomotora en una cierta representación, sino que busquemos conocer la estructura oculta bajo esa ley, es decir, querremos la ley sensomotora en su “expresión implícita”, que denotamos

$$S = \varphi(M, E)$$

Por lo anteriormente dicho, la expresión de la ley sensomotora implícita φ debe ser invariante a cambios de código. Por lo tanto, se expresará a través del conjunto de funciones que la dejan invariante, que denotamos $\Gamma(\varphi)$

$$\varphi \leftrightarrow \Gamma(\varphi)$$

donde $\Gamma(\varphi)$ retiene exactamente todo lo que es independiente del código en las diferentes representaciones que puede tomar φ . Dada una ley φ , el conjunto de funciones $\Gamma(\varphi)$ se define matemáticamente de la siguiente manera. Si denotamos

$$Sym(X) \stackrel{def}{=} \{f : X \rightarrow X, f \text{ invertible}\}$$

al conjunto de funciones invertibles de un conjunto X sobre sí mismo, entonces el conjunto de funciones $\Gamma(\varphi)$ puede definirse como

$$\Gamma(\varphi) \stackrel{def}{=} \{f \in \text{Sym}(M \times E) \text{ tal que } \varphi \circ f = \varphi\}$$

Supongamos dos agentes con la misma ley sensomotora implícita φ pero expresada en términos diferentes para cada agente (φ_1 y φ_2). Calcular el conjunto de funciones $\Gamma(\varphi)$ consiste en por encontrar las funciones f que permitan conectar ambas representaciones:

$$\varphi_1 = f \circ \varphi_2$$

¿Pero cómo obtener este conjunto de funciones con un sólo agente que siga una ley sensomotora φ ?

Podemos encontrar este conjunto de funciones f si disponemos de configuraciones agente-entorno diferentes,

$$(M1, E1), (M2, E2)$$

que proporcionen la misma S

$$\varphi(M1, E1) = \varphi(M2, E2) = S$$

y calculamos aquellas funciones que cumplan,

$$f(M1, E1) = (M2, E2)$$

Es decir, el conjunto $\Gamma(\varphi)$ que identifica la ley sensomotora vendrá expresado en términos del conjunto de funciones f tales

$$f : M \times E \rightarrow M \times E$$

que provocan modificaciones motoras y cambios en la configuración del entorno con la propiedad de que no modifican el valor sensorial S que recibe el agente. Matemáticamente, decimos que las funciones f son “funciones de simetría $M \times E$ con respecto a S ”.

Estas transformaciones son precisamente las que se corresponden con la definición de *movimientos compensables* (ver sección 3.2.3) que nos permitirán acceder a la estructura del espacio sensomotor del agente (definido en la sección 3.2.3).

Movimientos rígidos

Supondremos que todos los cambios (M, E) recogidos por el sistema son *rígidos*. Esto es lo mismo que suponer que no se producirán interferencias entre el agente y el entorno (como por ejemplo, cambios provocados en el entorno si el agente impacta sobre él). Matemáticamente, se traduce en que las funciones f de simetría pueden “desacoplarse”:

$$\begin{aligned} f : M \times E &\rightarrow M \times E \\ (m, e) &\rightarrow f(m, e) \end{aligned}$$

puede expresar como

$$\begin{aligned} f : M \times E &\rightarrow M \times E \\ (m, e) &\rightarrow (f(m), f(e)) \end{aligned}$$

Sólo queremos conservar los cambios que informan de las propiedades “objetivas” del espacio sensomotor (y aquellas que provienen de interferencias constituyen “ruido” en el problema). Por tanto, del conjunto de funciones f que pertenecen a $\Gamma(\varphi)$ quiero quedarme exclusivamente con un subconjunto: las que corresponden a movimientos rígidos, que denoto como $\Gamma^d(\varphi)$.

En resumen, tenemos un agente simple moviéndose por un entorno y recibiendo información sensorial suavemente. Suponemos que entorno y agente no se interfieren en sus movimientos y

que tras un tiempo tiene un conjunto de relaciones (S, M, E) . Matemáticamente consiste en asumir que en nuestro sistema se cumplen condiciones de diferenciabilidad (esto es, S , M , y E son variedades continuas, conexas, sobre las que operan funciones diferenciables), y que tenemos un agente holonómico ($P \leftrightarrow M$) que registra información sensorial S . Conocer cómo percibe un agente con ley sensomotora φ es conocer la estructura de su espacio sensomotor. Y esto pasa por encontrar $\Gamma^d(\varphi)$, esto es, el grupo de transformaciones rígidas que son simétricas para S . Precisamente este conjunto de funciones $\Gamma^d(\varphi)$ son el total de movimientos compensables del sistema agente-entorno.

3.3.2. Campo de vectores fundamentales sobre la variedad S

Uno de los conceptos esenciales en teoría de variedades diferenciales es la noción de “campo de vectores fundamentales”. En nuestro problema, un campo de vectores fundamentales sobre S , que denotaremos X^S , básicamente es otro modo de representar la variedad S en términos de vectores de sus espacios tangentes en cada uno de sus puntos. Para fijar ideas, supongamos que nuestra variedad S es una esfera. Calculemos los planos tangentes en cada uno de los puntos que constituyen la envolvente de S . En cada plano, definimos una base y sobre ella representamos un vector. El conjunto de vectores así obtenido constituye el campo de vectores fundamentales X^S de la variedad S . Dada S , el campo vector X^S será la herramienta matemática que nos permitirá conocer la estructura $\Gamma^d(\varphi)$.

La siguiente sección explica un método operativo para poder obtener un campo de vectores X^S sobre la variedad S que percibe el agente.

Espacio tangente y mapeo exponencial sobre una variedad

“Medir” en una variedad W es medir sobre su “arco de curva”, es decir, calcular longitudes a lo largo de su superficie. Si las variedades son diferenciables, existe un plano tangente en cada punto x de la variedad $T_x W$, y una función (mapeo exponencial, que denotamos $\exp(T_x W)$) que nos permite relacionar cada punto del plano con puntos de W en un entorno de x (ver Fig. 3.4).

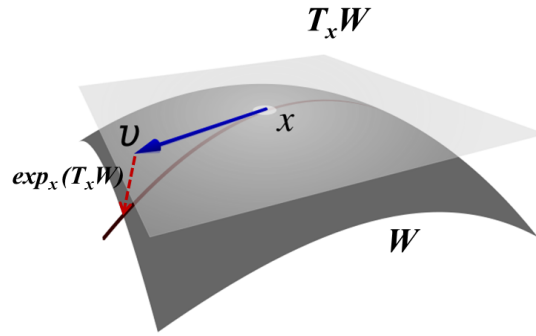


Figura 3.4: Ejemplo de espacio tangente y mapeo exponencial.

Supondremos que, al menos, en una parte del entorno (que denotamos V), si el agente está moviéndose de cierto modo (movimiento U), esto no “oscurece” la observación de los cambios que el entorno genera. Es decir, que aunque el agente se mueva, todos los cambios de este sub-entorno son registrados como cambios de S en los sensores. De otro modo, “siempre que se mueve el entorno, el agente se da cuenta sensorialmente”.

Matemáticamente, existe un conjunto de movimientos $U \subset M$ y existe un sub-entorno $V \subset E$, de modo que todos los cambios el entorno $\varphi(U, :)$ generan una función inyectiva entre V y S

$$\varphi(U, :) : V \rightarrow S$$

En condiciones de diferenciabilidad, sabemos que el conjunto de valores sensoriales registrados en esta situación

$$S = \varphi(U, V)$$

constituyen una variedad y $\varphi(U, :)$ es un difeomorfismo y, por tanto, tiene inversa que es diferenciable también.

Supongamos que el agente es capaz de moverse de un modo muy particular (no sabrá cómo hacerlo, pero nos interesa suponer que sí es posible para estimar qué efectos tendría si pudiera). Supondremos que el agente conoce todas las transformaciones rígidas $f \in \Gamma^d(\varphi)$ y por tanto, puede moverse en los términos motores de esas transformaciones. Recordemos que las transformaciones rígidas son funciones del producto cartesiano de M y E del tipo:

$$\begin{aligned} f : M \times E &\rightarrow M \times E \\ (m, e) &\rightarrow (f(m), f(e)) \end{aligned}$$

y que cumplen la propiedad de que los cambios motores $f(m)$ y los del entorno $f(e)$ se compensan entre sí. El agente, si como hemos supuesto conociese f , sabría qué movimiento motor realizar $f(m)$ aunque dicho movimiento no se vería necesariamente compensado por los cambios del entorno (es decir, los cambios en E no se garantizan que vengan dados por $f(e)$). Por lo tanto, el agente notaría cambios en sus sensores los cuales no podría percibir si el movimiento se compensara.

Estos cambios que percibe son los correspondientes a los movimientos compensables y los percibe, precisamente, porque en las condiciones supuestas no han sido compensados por el entorno.

Pero el hecho de que pueda conocer estos cambios en S supone un resultado muy relevante. Recordemos que nos interesa, dada una variedad, conocer su campo vector X^S porque será la herramienta matemática que nos permita conocer la estructura $\Gamma^d(\varphi)$.

A continuación mostraremos cómo a partir de los movimientos particulares $f(m)$ del agente se podrían generar con ellos los correspondientes cambios en S y obtener su campo vector asociado X^S . En una sección posterior, pasaremos a explicar cómo a partir de X^S podremos obtener la estructura de $\Gamma^d(\varphi)$ (véase sección 3.3.4).

Una vez generados los cambios en los valores de S , construimos X^S . Tal como se ha comentado inicialmente, un campo de vectores fundamentales se construye asociando a cada punto $p \in S$, un vector v_p del campo tangente en ese punto, $T_p S$. Pero nuestro modo de construir X^S no será a través de un algoritmo que genere la correspondencia $v_p \leftrightarrow p$ directamente sobre S , sino a través de movimientos en la variedad motora U . En concreto:

(i) definimos un único vector generador V_m en su respectivo espacio tangente $T_m U$ de la forma,

$$V_m = \frac{d}{dt}(f(e^{-tU}, m)|_{t=0}) \in T_m U$$

donde $f(e^{-tU}, m)$ se refiere al mapeo exponencial que nos permite formalizar movimientos pequeños sobre la variedad U en un entorno del punto $m \in U$, a lo largo del arco de curva de la propia variedad, a partir del movimiento rígido $f \in \Gamma^d(\varphi)$ sobre el espacio tangente $T_m U$. Asumimos que estas variaciones $(\frac{d}{dt})$ son pequeñas puesto que tienen lugar en el espacio tangente $T_m U$.

que (ii) permite definir una orden motora generadora $M_U(t)$, definida como:

$$\begin{cases} M_U(0) = m \\ \dot{M}_U(0) = -\frac{d}{dt}(f(e^{-tU}, m)|_{t=0}) \end{cases}$$

(iii) que es la base para construir un sistema generador de X^S a través de la ley sensomotora φ ,

$$X^S = \frac{d}{dt}\varphi(M_U(t), V)$$

Con este procedimiento tendríamos definido un campo de vectores generadores de X^S que nos permitiría conocer la estructura de $\Gamma^d(\varphi)$ (ver sección 3.3.4). Pero recordemos que este procedimiento no podría seguirlo el agente puesto que no conoce las funciones $f \in \Gamma^d(\varphi)$ necesarias para poder ejecutarlo (conocerlas implicaría conocer la estructura de $\Gamma^d(\varphi)$, que es precisamente la que queremos obtener).

A continuación, por tanto, explicaremos estas dos cuestiones: (i) en la sección 3.3.3, mostraremos cómo el agente podría obtener órdenes motoras generadoras a partir de sus movimientos y (ii) finalmente, cómo acceder a la estructura $\Gamma^d(\varphi)$ a través del campo de vectores X^S .

3.3.3. Órdenes motoras generadoras en X^S

El método operativo para trabajar con variedades se reduce a trabajar en sus espacios tangentes. Por ello, supondremos que tanto los cambios en la configuración del agente como del entorno son muy pequeños (diferenciales) y, por lo tanto, las consecuencias sensoriales de la interacción del agente con su entorno también quedarán registradas sobre su espacio tangente.

El conjunto de todos los estímulos recibidos por el agente a través de su interacción sensomotora con el mundo (medidos en su espacio tangente $S = T\varphi(M, E)$), puede dividirse en dos subconjuntos: (i) uno correspondiente a las variaciones sensoras debidas solo al movimiento del agente ($S = T\varphi(M, E_0)$), y (ii) otro correspondiente a las variaciones sensoras debidas solo al movimiento del entorno ($S = T\varphi(M_0, E)$).

- *Si el cuerpo del agente se mueve dejando el entorno fijo:* es posible acceder a una base del espacio tangente $S = T\varphi(M, E_0)$ a partir de las relaciones (correspondencias sensomotoras) entre los movimientos del agente (moviéndose en su variedad M) y las correspondientes respuestas sensoras S obtenidas (medidas sobre el propio espacio tangente) (ver Fig. 3.5a). Obtener una base del espacio tangente $S = T\varphi(M, E_0)$ es, en nuestro problema, acceder a la representación del estado del cuerpo (definida en la sección 3.2.2) y conocer su dimensión es, a su vez, conocer el mínimo número de parámetros en que puede caracterizarse el cuerpo “percibido”.
- *Si, por el contrario, el agente deja su cuerpo inmóvil y registra los cambios sensoriales provocados por los cambios del entorno:* es posible acceder a una base del espacio $S = T\varphi(M_0, E)$ y por tanto a la representación del estado del entorno y a su dimensión (ver Fig. 3.5b). Esto es posible porque estamos asumiendo que el agente es capaz de percibir todos los cambios del entorno (siempre que actúe en el sub-entorno V definido anteriormente).

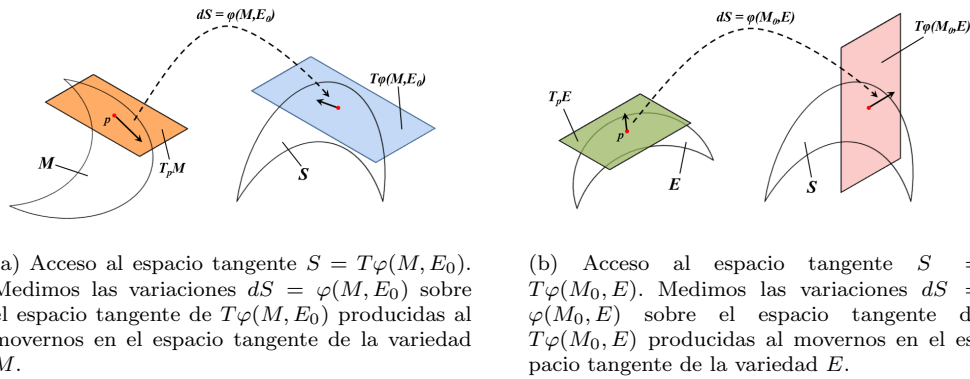


Figura 3.5: Extracción de espacios tangentes $T\varphi(M_0, E)$ y $T\varphi(M, E_0)$

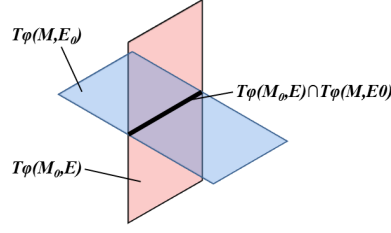


Figura 3.6: Espacio tangente intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$.

El espacio tangente intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ representa el conjunto de todos los estímulos recibidos por el agente que pueden generarse por una combinación de movimientos del cuerpo y del entorno (donde ninguno es nulo). Esto quiere decir que *todos los posibles movimientos compensables agente-entorno tienen respuesta sensorial en dicha intersección*, puesto que dado cualquier movimiento del agente, existe un movimiento del entorno que puede dejar invariante la respuesta sensorial (ver Fig. 3.6).

El problema que teníamos para construir el campo de vectores fundamentales de X^S es que el agente no puede construir órdenes motoras generadoras $M_U(t)$, puesto que no conoce las funciones $f \in \Gamma^d(\varphi)$ necesarias para generar $M_U(t)$. Sin embargo, sabemos que el espacio intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ contiene el conjunto de las respuestas sensoriales S asociadas a los movimientos compensables.

Podemos asegurar que, si el agente es capaz de generar una orden motora $M(t)$ cualquiera (que no tiene por qué definirse a partir de funciones $f \in \Gamma^d(\varphi)$), pero cuya respuesta sensorial S , mediada por la ley sensomotora φ , se encuentra en dicha intersección,

$$M(t) \in T\varphi(M_0, E) \cap T\varphi(M, E_0)$$

entonces esa orden motora $M(t)$ será equivalente a una orden motora generadora $M_U(t)$, debido a que las órdenes motoras generadoras se construyen a partir de una transformación rígida $f \in \Gamma^d(\varphi)$ que, por definición (ver sección 3.3.1) dejan invariante la S .

Por lo tanto, *cualquier orden motora que tenga una respuesta sensorial* (a través de la ley sensomotora) *en el espacio tangente intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$, será una orden motora generadora de X^S .*

$$\frac{d}{dt}\varphi(M(t), V) = \frac{d}{dt}\varphi(M_U(t), V) \Rightarrow \frac{d}{dt}\varphi(M(t), V) = X^S$$

En el capítulo siguiente se mostrará un procedimiento mediante el cual se pueden obtener órdenes motoras generadoras a partir de una base del espacio $T\varphi(M_0, E) \cap T\varphi(M, E_0)$. El agente dispondrá así de una alternativa para generar X^S sin necesidad de conocer las funciones $f \in \Gamma^d(\varphi)$.

3.3.4. Estructura de $\Gamma^d(\varphi)$ a partir de X^S

La última parte de este capítulo, tal como hemos indicado, la dedicaremos a explicar cómo puede conocerse la estructura del espacio sensomotor $\Gamma^d(\varphi)$ partir del campo de vectores fundamentales X^S . Para ello, utilizaremos nociones de grupos y álgebras de Lie que pasamos a indicar.

Grupos y álgebras de Lie

Decimos que un conjunto de elementos C , junto con una operación binaria \circ , tiene estructura matemática de grupo G si la operación \circ es cerrada en G y cumple la propiedad asociativa, y además existe elemento neutro y, para cada elemento de C existe su elemento inverso. Por tanto,

conocer la estructura de un grupo G es conocer las propiedades que se cumplen cuando actúa la operación de grupo \circ sobre sus elementos $g_1, g_2 \in G$, esto es $g_1 \circ g_2$.

Un conjunto de transformaciones con estructura de grupo G sobre una variedad W es de Lie si sus transformaciones son continuas y diferenciables, es decir, que admite una representación en términos de su espacio tangente. Y el espacio tangente de un grupo de Lie L es un álgebra de Lie si existe una operación en él conocida como corchete de Lie $[\cdot, \cdot]$ que, a partir de dos elementos de $X, Y \in L$, cuantifica cómo de lejos está “multiplicarlos en una dirección o en otra”. Esto es:

$$[X, Y] = XY - YX$$

El corchete de Lie debe satisfacer las siguientes propiedades: (i) cumple la identidad de Jacobi, (ii) es bilineal, y (iii) la aplicación del corchete sobre un único vector $[X, X]$ es nula.

Consideremos el siguiente ejemplo: Si visualizamos el grupo G como operaciones con matrices (donde su operación de grupo es la composición de matrices), las propiedades de la operación sobre $f, g \in G$

$$f \circ g(\cdot) = f(g(\cdot))$$

quedan, de algún modo, recogidas en el resultado de aplicar la operación propia de L entre los elementos X, Y de la correspondiente álgebra L de G (en nuestro ejemplo, la operación suma entre vectores) que se corresponden con las transformaciones $f, g \in G$

$$X + Y$$

En este ejemplo, el corchete de Lie es el modo de cuantificar el grado de no-conmutatividad (en los casos en los que el corchete de Lie sea no nulo) del grupo G de partida. De algún modo, el corchete de Lie al aplicarse sobre los vectores del álgebra $X, Y \in L$ que representan operaciones del grupo G , y observar su comportamiento, $[X, Y]$, es capaz de rescatar propiedades sobre la operación de composición del grupo. De otra forma, el corchete de Lie en L es capaz de proporcionarnos información sobre la estructura del grupo G asociado. En general, la estructura de un grupo G de Lie puede recuperarse sólo parcialmente a partir de su álgebra de Lie (la condición directa, de grupo a álgebra, es siempre cierta) pero no la contraria, aunque en ciertas condiciones matemáticas, puede recuperarse de manera completa (ver Anexo A).

En las condiciones de nuestro problema, puede probarse que:

1. $\Gamma^d(\varphi)$ tiene estructura de grupo.
2. El grupo $G = \Gamma^d(\varphi)$ es de Lie (es decir, sus transformaciones pueden ser expresadas como composiciones de transformaciones diferenciables)
3. A partir de campos de vectores fundamentales X^S de S puede obtenerse la estructura del álgebra de Lie L del grupo $\Gamma^d(\varphi)$, del siguiente modo:

Sean dos transformaciones $f, g \in \Gamma^d(\varphi)$ que inducen dos campos vectoriales fundamentales X^S, Y^S sobre S , y a su vez, se relacionan con dos elementos X, Y de su correspondiente álgebra de Lie. Puede demostrarse que X^S, Y^S está relacionada con X, Y a través del corchete de Lie de L , del siguiente modo:

$$[X^S, Y^S] = [X, Y]$$

4. La estructura del álgebra de Lie L de $\Gamma^d(\varphi)$ codifica la estructura de $\Gamma^d(\varphi)$.

Pero para poder aplicar correctamente este resultado en nuestro problema, hay que garantizar que diferentes transformaciones $f, g \in \Gamma^d(\varphi)$ generan diferentes campos vectoriales fundamentales X^S, Y^S . En caso contrario, tendríamos dos transformaciones f, g que generarían el mismo efecto X^S , y nos llevarían a un resultado nulo, $[X^S, X^S] = 0$, por las propiedades del corchete de Lie.

Necesitamos garantizar que esto no ocurre para poder aplicar el procedimiento antes indicado.

Pero como el agente no puede garantizar transformaciones con estas propiedades (porque no conoce $\Gamma^d(\varphi)$) sólo podremos conocer la estructura de $\Gamma^d(\varphi)$ a través de su álgebra de Lie en un subconjunto $\Gamma_0^d(\varphi) \subset \Gamma^d(\varphi)$ donde todas las transformaciones generen cambios sensoriales S diferentes. Hemos demostrado que existe un método matemático para extraer la estructura de un subconjunto $\Gamma_0^d(\varphi) \subset \Gamma^d(\varphi)$ a través del campo de vectores fundamental X^S obtenido a partir de las relaciones obtenidas por un agente sensomotor entre sus movimientos y las respuestas sensoriales que obtiene en respuesta.

Conclusiones

Hemos, por tanto, mostrado que según los trabajos [10, 11], para un organismo que interactúa sensomotormente con su entorno, procesando las relaciones sensoriales que recibe en respuesta a sus movimientos, sin tener ninguna información a priori sobre el espacio físico, y sin conocer la correspondencia entre su cuerpo y su sistema nervioso (véase sección 3.2.1), *existe un procedimiento matemático* mediante el cual puede llegar a: (i) distinguir cuerpo de entorno “percibidos” y acceder a sus estructuras, y (ii) acceder a la estructura común cuerpo-entorno que constituye su *espacio sensomotor* (ver sección 3.2.3). Es decir que *tanto los límites del cuerpo como la geometría y dimensionalidad del espacio son deducibles, sin ningún conocimiento a priori, a partir de los vínculos entre sus estímulos sensoriales y sus movimientos motores.*

3.4. Algoritmo

En esta sección se mostrará el método algorítmico que, basado en el modelo explicado, simulará estar implementado en el cerebro de un organismo simple y que haciendo uso de los conceptos matemáticos descritos en la sección anterior, permitirá: (i) extraer las bases de los espacios tangentes $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ (esto es, las estructuras de representación de los estados del cuerpo y el entorno), (ii) extraer una base del espacio tangente $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ (relacionada con los movimientos compensables) y (iii) extraer una base de vectores generadores V_m con los que generar una orden motora generadora $M_U(t)$ del campo de vectores fundamentales X^S asociado a S .

El algoritmo se divide en 5 etapas:

1. **Condiciones iniciales:** Se dispone de un agente sensomotor en un entorno al cual tiene total acceso y con el que puede interactuar sensomotormente. El agente parte de una posición inicial M_0 y el entorno tiene una configuración inicial E_0 .
2. **Obtención de una base de $T\varphi(M, E_0)$:** El agente se mueve diferencialmente en entornos locales de M_0 (esto es, en términos de geometría diferencial, moverse en el espacio tangente de M en el punto M_0 a través de un vector dM) y obtiene pequeñas variaciones de S (esto es, medidas en el espacio tangente $S = T\varphi(M, E_0)$ en puntos muy próximos a S_0).

$$S_0 + dS = \varphi(M_0 + dM, E_0)$$

Como las medidas, tanto de M como de S , se determinan en espacios tangentes con estructura de espacio vectorial, todo el aparato matemático que necesitamos se reduce al álgebra lineal. Se utilizan las pequeñas variaciones de S para determinar (i): por regresión lineal, un sistema generador del espacio tangente $S = T\varphi(M, E_0)$ de dimensión k , y (ii), a partir del sistema generador, se reduce la dimensión k mediante el método de descomposición de valores singulares (SVD) para obtener una base del espacio tangente $T\varphi(M, E_0)$ y, por tanto, su dimensión $\dim_m < k$.

Obtenemos, como resultado, una base de \dim_m vectores de dimensión $n = \dim(S)$ que expresaremos en forma de matriz de \dim_m columnas de y n filas.

3. **Obtención de una base de $T\varphi(M_0, E)$:** El agente se mantiene estático mientras el entorno se mueve y se obtiene una base de $T\varphi(M_0, E)$. Siguiendo el mismo método del punto 2 pero con movimientos diferenciales del entorno en vez del agente, se obtendrá una base de $T\varphi(M_0, E)$ y su dimensión (dim_e). Es decir, como resultado de esta fase, tendremos una base de dim_e vectores de dimensión $n = dim(S)$ que expresaremos en forma de matriz de dim_e columnas de y n filas.
4. **Obtención de una base de $T\varphi(M_0, E) \cap T\varphi(M, E_0)$:** Concatenando las bases $T\varphi(M_0, E)$ y $T\varphi(M, E_0)$ (obtenidas en las etapas 2 y 3) se obtiene un sistema generador del espacio tangente suma $T\varphi(M_0, E) + T\varphi(M, E_0)$. A través de la descomposición de valores singulares, se puede obtener tanto una base del conjunto suma como su dimensión (dim_{suma}). Además, el método SVD (ver Anexo B) nos permite acceder al espacio nulo del conjunto suma, que es por definición el espacio tangente intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$.

Se obtiene entonces tanto la base del espacio intersección como su dimensión (dim_{int}), que será

$$dim_{int} = dim_m + dim_e - dim_{suma}$$

esto es, obtenemos, como resultado, una base de dim_{int} vectores de dimensión $n = dim(S)$ que expresaremos en forma de matriz de dim_{int} columnas de y n filas.

5. **Obtención de una base de vectores generadores de X^S :** Toda orden motora que tenga consecuencias sensoriales en la intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ es una orden motora generadora de X^S (ver sección 3.3.3). Como se dispone de una matriz de la base de $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ (fase 4), una matriz de la base de $T\varphi(M, E_0)$ (fase 2) y del vector dM con el que se generó $T\varphi(M, E_0)$ (fase 2), entonces podremos recorrer el camino inverso y utilizar las matrices pseudoinversas de las obtenidas para llegar desde el espacio tangente intersección a un conjunto de vectores $\{V_m\}$, que como se vio en la sección 3.3.3, son un conjunto de vectores generadores de X^S (a través de una orden motora $M_U(t)$),

$$\{V_m\} = dM \cdot T\varphi^{-1}(M, E_0) \bullet [T\varphi(M_0, E) \cap T\varphi(M, E_0)]$$

Este conjunto $\{V_m\}$ con el que se generará la orden motora $M_U(t)$ y, con ello, el campo de vectores X^S será un conjunto de dim_{int} vectores de dimensión dim_m (recordemos que es un conjunto generador motor) que expresaremos en forma de matriz de dim_{int} columnas de y dim_m filas.

Capítulo 4

Entorno experimental

En este capítulo mostraremos la validez del modelo propuesto y del algoritmo planteado en secciones previas. Para ello diseñamos un entorno experimental hacemos dos tipos de experimentos: (i) el primero discurrirá en el entorno de simulación donde no tendremos restricciones en los grados de libertad de los dispositivos sensomotores del sistema analizado. Este experimento nos permitirá mostrar la potencia del método y no se preocupará ni de aspectos relacionados con la implementación física del sistema, ni de la complejidad del entorno. (ii) el segundo aborda un caso real (un robot *Lego* con un dispositivo visual y dos motores) moviéndose en un entorno controlado.

4.1. Diseño del entorno virtual y del agente sintético

En el marco experimental desarrollado se introducen los siguientes elementos:

Entorno

Espacio formado por seis elementos emisores de estímulos sensoriales situados aleatoriamente, próximos unos a otros. Los seis elementos emiten estímulos de tipo luminoso y cuatro de ellos, además, emiten estímulos de tipo sonoro. Los otros dos funcionan como generadores de perturbaciones en el medio que se detectan de manera táctil. La intensidad de las tres fuentes de estímulo sigue una relación inversamente proporcional al cuadrado de la distancia.

Agente

El agente simulado (ver Fig. 4.1) que representa un organismo simple presenta (1) dispositivos de interacción sensorial con el medio (sistema sensor), (2) un sistema vestibular para su movimiento (sistema motor), y (3) un sistema nervioso primitivo interno (red de conexiones nerviosas).

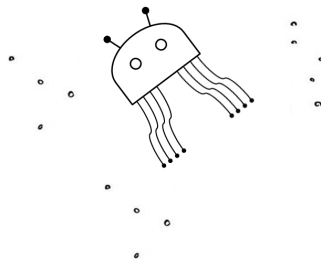


Figura 4.1: Agente simulado con tres tipos de dispositivos sensomotores.

4.1.1. Sistema sensor del agente

El organismo se encuentra situado en el entorno descrito anteriormente. Sus mecanismos sensores son de tres tipos: (i) dos dispositivos visuales, (ii) dos dispositivos auditivos y (iii) ocho dispositivos táctiles.

A) DISPOSITIVO VISUAL

Estructura

El dispositivo visual del agente está constituido por dos “ojos” situados a la misma altura, con posibilidad de movimientos independientes pero orientados en una dirección común (como el sistema visual humano aunque no necesariamente con la complejidad del ojo de un organismo real), cada uno de los cuales está formado por 40 células fotosensibles distribuidas aleatoriamente en una retina bidimensional.

Además, existen dos dispositivos independientes por cada ojo (dispositivos “obturadores” similares a los de una cámara) que regulan la cantidad de luz que accede a la retina.

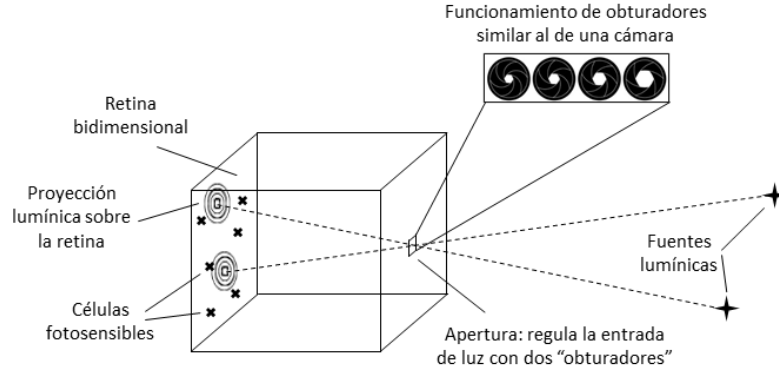


Figura 4.2: Ejemplo esquemático del funcionamiento del dispositivo visual. La proyección de las fuentes del entorno se modelan siguiendo el modelo de cámara estenopeica.

Intensidad lumínica sensorial

Para simular la interacción de la luz con el ojo, se calcula la proyección de cada fuente luminosa sobre la retina siguiendo el modelo de la cámara estenopeica suponiendo una distancia focal unitaria (ver Fig. 4.2). El siguiente paso determina la cantidad de luz que, a partir de esa proyección, le llega a cada sensor de la retina.

Dado un sensor con coordenadas en la retina (x, y) en un rango $[0, 1]$, y la proyección de una fuente luminosa i centrada en la posición (x_0, y_0) de la retina y a distancia d del foco luminoso, la intensidad lumínica I_i^{lum} que recibe dicho sensor proveniente de la fuente i se calcula (ver Fig. 4.3):

$$I_i^{lum} = e^{-\frac{(x_i - x_0)^2 + (y_i - y_0)^2}{d}}$$

Ley perceptual del agente

La estimulación I^{lum} total que recibe cada sensor de la retina (y que define el modo de funcionamiento del dispositivo visual) es la suma de las intensidad lumínicas de cada fuente (ver apartado previo) reguladas por la acción de los “obturadores” que actúan de filtro del flujo luminoso.

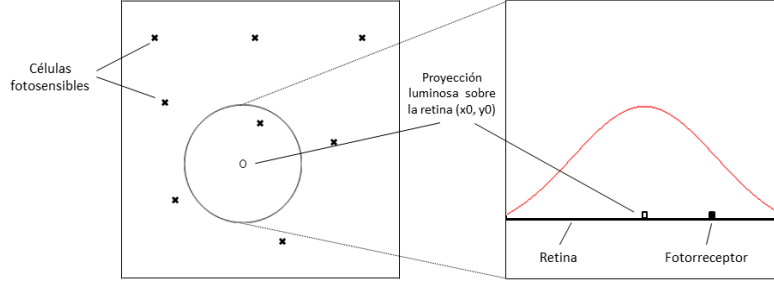


Figura 4.3: Ejemplo esquemático del funcionamiento de la retina. El total de intensidad luminosa que recibe el fotorreceptor se modela mediante una gaussiana en función de la distancia entre el fotorreceptor y la proyección luminosa sobre la retina.

$$I^{lum} = (d_1 + d_2) \cdot (I_1^{lum} + I_2^{lum} + \dots + I_n^{lum})$$

siendo I_i^{lum} la intensidad recibida por el sensor por parte de la fuente i , n el número total de fuentes (en nuestro entorno $n = 6$), y d_1 y d_2 las aperturas de los dos mecanismos obturadores.

B) DISPOSITIVO AUDITIVO

Estructura

El dispositivo auditivo del agente está constituido por dos “antenas” situados de manera rígida a ambos lados del organismo, a la misma altura y de manera simétrica.

Ley psicofísica

Ambas “antenas” son sensibles a la intensidad sonora, con un perfil de sensibilidad que favorece las fuentes auditivas situadas frente a los dispositivos sensoriales (penalizando los cambios de orientación entre fuente y receptores) (ver Fig. 4.4). La intensidad sonora I_i^{son} que recibe cada antena proveniente de la fuente auditiva i es:

$$I_i^{son} = \frac{1}{d} \cdot \frac{(z - z_i)}{\sqrt{(y - y_i)^2 + (z - z_i)^2}} \cdot \frac{(z - z_i)}{\sqrt{(x - x_i)^2 + (z - z_i)^2}}$$

siendo (x, y, z) las coordenadas de la antena, (x_i, y_i, z_i) las coordenadas de la fuente sonora i y siendo d la distancia de la fuente i a la antena.

La intensidad total sonora I^{son} que recibirá cada antena será la suma de las intensidades proporcionadas por cada fuente sonora:

$$I^{son} = I_1^{son} + I_2^{son} + \dots + I_n^{son}$$

siendo I_i^{son} la intensidad recibida por la antena por parte de la fuente i y n el número total de fuentes sonoras (en nuestro entorno $n = 4$).

C) DISPOSITIVO TÁCTIL

Estructura

El dispositivo táctil está formado por ocho receptáculos (“flagelos”) situados a ambos lados del organismo de manera simétrica.

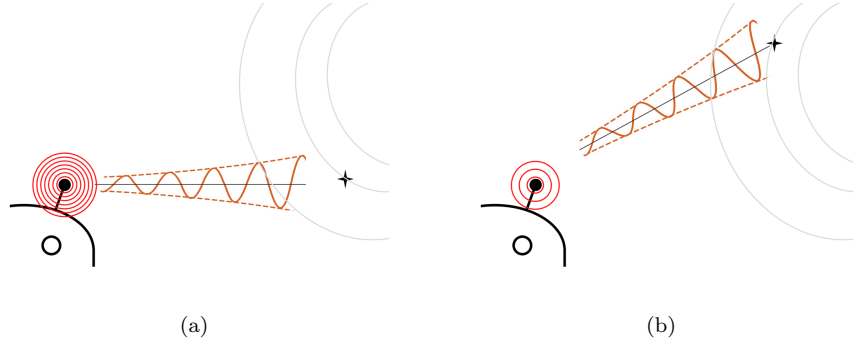


Figura 4.4: Ejemplo esquemático del funcionamiento del dispositivo auditivo. La antena favorece (a) las fuentes auditivas situadas frente a los dispositivos sensoriales y penaliza (b) los cambios de orientación entre fuente y antena.

Ley psicofísica

Los dispositivos se ven afectados por perturbaciones del medio, mediante el desplazamiento relativo a su posición de reposo (ver Fig. 4.5). Ese desplazamiento es dependiente de la distancia del dispositivo con respecto a la fuente generadora de la perturbación a la que el sistema es sensible. Supondremos que existe un fenómeno de saturación que es modelado como una relación $x^{1/2}$.

La intensidad total táctil I^{tac} que recibirá cada flagelo será la suma de las intensidades proporcionadas por cada fuente generadora de perturbaciones:

$$I^{tac} = \sqrt{d_1} + \sqrt{d_2} + \dots + \sqrt{d_n}$$

siendo d_i la distancia del flagelo a la fuente i y n el número total de fuentes generadoras de perturbaciones (en nuestro entorno $n = 4$).

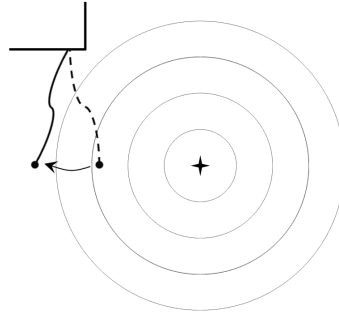


Figura 4.5: Ejemplo esquemático del funcionamiento del dispositivo auditivo. El agente percibe el desplazamiento de su flagelo debido a las perturbaciones provocadas por la fuente.

La estructura del espacio sensor, por tanto, estará compuesta por un total de 90 sensores (40 células fotosensibles en cada retina, dos sensores auditivos, y ocho sensores táctiles).

4.1.2. Sistema motor del agente

Por sistema motor no nos referimos ni a la estructura mecánica de locomoción del agente ni a su sistema fisiológico (que llamaremos en conjunto “sistema efector-motor”) sino a la estructura de posibles configuraciones del agente dentro de su entorno (a partir de sus posibles cambios de

posición y orientación del cuerpo, de la orientación de sus dispositivos sensoriales, etc.) y al número de grados de libertad en que ésta se define. A este espacio de configuración es al que llamaremos sistema motor.

Para fijar las ideas, supongamos un avión. El “sistema motor” al que nos hemos referido quedaría definido en términos de la estructura de sus movimientos (traslaciones en las tres dimensiones del espacio, y las tres rotaciones en torno a los ejes de Euler) y del número de grados de libertad de los mismos (6). No debe confundirse con el número de dimensiones (3) del espacio euclídeo en el que se está moviendo ni con la estructura mecánica de sus dispositivos de movimiento (motores, dispositivos de dirección, etc.). El sistema motor de nuestro agente, tal como ha sido definido en la sección previa, tendría la siguiente estructura: (i) tres parámetros para fijar la posición del cuerpo, (ii) tres para determinar la orientación del cuerpo, (iii) seis para determinar la orientación de ambos “ojos” (tres por cada ojo), (iv) y cuatro para determinar la apertura de los obturadores de ambos ojos (dos por cada ojo). Tanto el sistema auditivo como el táctil están situados de manera rígida en el cuerpo y, por tanto, no introducen ningún grado de libertad adicional al sistema motor. El sistema motor descrito es por tanto de dimensión dieciséis.

4.1.3. Sistema nervioso del agente

Supondremos un sistema nervioso básico formado por dos subsistemas: sistema neuronal sensor y sistema neuronal motor.

El *sistema nervioso sensor* está compuesto por 90 neuronas que reciben el estímulo de los 90 sensores totales del organismo. Para simular los procesos de plasticidad y auto organización neuronal que permitirán al agente ajustar su comportamiento al mundo, se supone que se encuentran distribuidas inicialmente en correspondencia biyectiva (simulando una especie de sistema nervioso en un estado embrionario de partida con una única conexión neurona-sensor) y que evolucionarán (supondremos que se reorganizarán mediante combinaciones lineales) según el agente interactúe sensomotormente con el medio (ver Fig. 4.6).

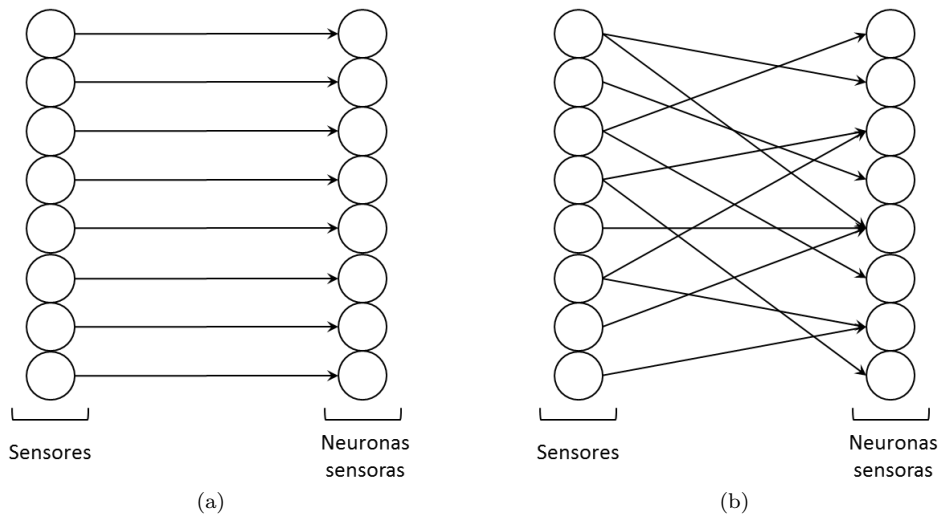


Figura 4.6: Ejemplo esquemático del sistema nervioso sensor antes (a) y después (b) de la auto organización neuronal.

El *sistema nervioso motor* está compuesto por 40 neuronas (número arbitrario para simular cierta complejidad neuronal motora) que se conectan con los elementos del “sistema efector-motor” (que sería el encargado de mover tanto el cuerpo como los dispositivos sensoriales) dando lugar a una configuración concreta en los 16 parámetros del “sistema motor” anteriormente definido. Esta

conexión está arbitrariamente distribuida simulando la complejidad neuronal motora. Al contrario que el sistema neuronal sensor, esta conexión es fija y no es adaptable, dejando la función de reorganización neuronal exclusivamente al sistema nervioso sensor, tal como hemos indicado previamente (ver Fig. 4.7).

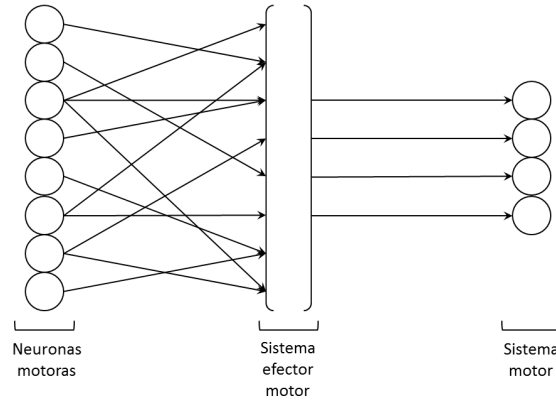


Figura 4.7: Ejemplo esquemático del sistema nervioso motor.

4.2. Experimento tipo 1: Entorno simulado

4.2.1. Aplicación del algoritmo

Tal como se indicó en el capítulo anterior, el modo en el que el sistema extrae regularidades de su entorno, medidas a partir de sus interacciones sensomotoras, se basa en *pequeños movimientos* aleatorios que le permiten obtener en respuesta *pequeñas variaciones sensoriales*. A partir de estas relaciones sensomotoras (cambios sensoriales producidos por cambios motores) el agente extrae la información sobre el espacio en el que se halla (las propiedades de dicho espacio están implícitas en la forma en la que cambia la información sensorial en respuesta al movimiento del agente).

La inclusión de tres tipos de sensores no sólo pretende simular un organismo con cierta complejidad sensorial. Además, permitirá testear si la extracción de regularidades del entorno mediadas por leyes sensomotoras es independiente del tipo de estímulo. Es decir, si la información que gestiona un organismo multimodal es independiente del tipo de sensores que posee y de las leyes físicas y psicofísicas involucradas.

Detalles de implementación

El algoritmo que se ejecuta para llevar a cabo el experimento es el definido en el capítulo anterior (ver sección 3.4). La mayor parte del algoritmo consiste en operaciones matemáticas relacionadas con el álgebra lineal y la geometría diferencial, que son independientes de las configuraciones de agente y entorno elegidas. Sin embargo, en las etapas 2 y 3 del algoritmo, tanto la ejecución movimientos “diferenciales” de agente y entorno, como la obtención de respuestas sensoriales, sí que dependen de la configuración particular del experimento. Por tanto, se van a explicar, a continuación, los detalles relativos a estas cuestiones de acuerdo a los parámetros concretos de agente y entorno definidos para este experimento. Una información más detallada sobre la implementación del algoritmo puede consultarse en el Anexo B.

Como todo el aparato matemático utilizado se reduce al álgebra lineal (ver sección 3.4), se utilizarán matrices tanto para definir las estructuras que determinan los sistemas nerviosos sensor y motor, como para modelar los movimientos de agente y entorno.

Estructuras matriciales

Las matrices que se usarán para representar tanto los sistemas nerviosos sensor y motor como el entorno, se instancian de la siguiente manera:

- **Matriz W_S .** Representa el sistema nervioso sensor y simula la conexión entre las 90 neuronas sensoras y los 90 sensores del organismo. Inicialmente será una matriz identidad (recordemos que el estado inicial del proceso de autoorganización sensorial considera una correspondencia biyectiva) de tamaño 90×90 .
- **Matriz W_M .** Representa la acción de las 40 neuronas motoras sobre los 16 parámetros del sistema motor del agente. Matemáticamente, tendremos una matriz de tamaño 16×40 y con valores aleatorios en el rango $[-1, 1]$.
- **Matriz E .** Representa las tres coordenadas de posición de cada uno de los 6 emisores de estímulos sensoriales del entorno. Matemáticamente, tendremos una matriz de tamaño 3×6 y con valores aleatorios en los siguientes rangos:

$$\text{Coordenadas } (x, y) \in [0, 50]; \text{ Coordenada } z \in [100, 101]$$

Movimientos diferenciales

Un movimiento motor M (que será un vector de 40 componentes, representando la acción de las 40 neuronas que lo generan) del agente, generará una configuración del sistema motor P (representado por un vector de 16 componentes, que definen su configuración motora) a través de su sistema nervioso motor W_M de la siguiente manera:

$$P = W_M \cdot M$$

Cuando el agente realiza movimientos diferenciales dM en torno a su posición inicial M_0 (fase 2 del algoritmo), matricialmente esto es lo mismo que hacer:

$$dP = W_M \cdot (M_0 + dM)$$

Para simular estos movimientos diferenciales dM en nuestro experimento, a partir de un movimiento aleatorio M , éste se divide en 9 movimientos diferenciales iguales dM_1, \dots, dM_9 (representados también por vectores de 40 elementos) tales que:

$$dM_1 + dM_2 + \dots + dM_9 = M$$

Obtención de la respuesta sensorial S

Para generar la respuesta sensorial S (representada por un vector de 90 componentes, una por cada sensor) asociada a dichos movimientos diferenciales (etapa 2 del algoritmo, ver sección 3.4), el agente realizará un conjunto de movimientos $\{M_i\}$, de forma que para cada movimiento particular M_i del conjunto, el agente ejecutará los 9 movimientos diferenciales asociados (tal cual se ha descrito anteriormente) generando la respuesta sensorial S_i asociada a cada movimiento, es decir:

$$\forall M_i \in \{M_i\} :$$

$$\begin{aligned} dS_1 &\leftarrow dP_1 = W_M \cdot (M_0 + dM_1) \\ dS_2 &\leftarrow dP_2 = W_M \cdot (M_0 + dM_2) \\ &\vdots \\ dS_9 &\leftarrow dP_9 = W_M \cdot (M_0 + dM_9) \end{aligned}$$

A partir de cada configuración particular dP_i del agente y de la configuración del entorno E , se obtiene su correspondiente respuesta sensorial $dS_i = T\varphi(dP_i, E)$ utilizando la ley psicofísica visual, auditiva o táctil (sección 4.1.1) correspondiente a cada tipo de sensor. Es decir, cada componente del vector dS_i , habrá sido calculada a partir de cada dP_i de acuerdo a la ley psicofísica asociada al sensor que representa esa componente.

En la configuración de nuestro experimento, las 80 primeras componentes de dS_i estarán asociadas a los sensores visuales, las 2 siguientes a los sensores auditivos, y las 8 últimas a los sensores táctiles, de manera que:

$$\forall dS_i \leftarrow dP_i = W_M \cdot (M_0 + dM_i) :$$

$$\begin{aligned} dS_i(1) &\leftarrow \text{ley psicofísica visual} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \\ &\vdots \\ dS_i(80) &\leftarrow \text{ley psicofísica visual} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \\ dS_i(81) &\leftarrow \text{ley psicofísica auditiva} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \\ dS_i(82) &\leftarrow \text{ley psicofísica auditiva} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \\ dS_i(83) &\leftarrow \text{ley psicofísica táctil} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \\ &\vdots \\ dS_i(90) &\leftarrow \text{ley psicofísica táctil} && \leftarrow dP_i = W_M \cdot (M_0 + dM_i) \end{aligned}$$

A partir de las respuestas sensoriales S obtenidas en los sensores a partir de cada movimiento M , se simula el paso de información de los sensores a las neuronas motoras mediante la matriz W_S ¹:

$$S' = W_S \cdot S$$

Una vez el cerebro del agente dispone de las respuestas sensoriales a sus movimientos S' , ya se dispone de todos los elementos para obtener la base del espacio tangente $T\varphi(M, E_0)$ siguiendo el método explicado en la sección 3.4 (y más en profundidad en el anexo B). La obtención de las respuestas sensoriales S debidas a cambios en el entorno E (etapa 3) se ha modelado exactamente con el mismo procedimiento descrito.

4.2.2. Resultados

Para poner a prueba la validez del algoritmo que implementa el modelo sensomotor propuesto, se han llevado a cabo una serie de simulaciones realizadas en el entorno de *MATLAB*. Obtuvimos los siguientes resultados:

Cálculo del $T\varphi(M, E_0)$

Al ejecutar la parte de la simulación correspondiente al movimiento del agente dejando el entorno fijo (etapa 2 del algoritmo) se obtuvieron los siguientes resultados:

La figura muestra la evolución de la estimación del ratio entre dos valores singulares sucesivos de la matriz que codifica la base de $T\varphi(M, E_0)$ (ver Fig. 4.8). El máximo de ese ratio muestra la frontera entre los valores singulares nulos y no nulos de la matriz analizada (ver anexo B). Los valores singulares no nulos revelan la dimensión del espacio tangente de los inputs sensoriales cuando el entorno está inmóvil ($T\varphi(M, E_0)$). El algoritmo descubre así que hacen falta 14 parámetros para caracterizar la configuración del cuerpo de los 16 utilizados, debido a que los efectos de los dispositivos obturadores se combinan en cada ojo en un solo parámetro, que es la modificación del flujo de luminosidad total que afecta a cada retina.

¹Para más información sobre el proceso de autoorganización sensorial de W_S descrito, consultar el Anexo B.

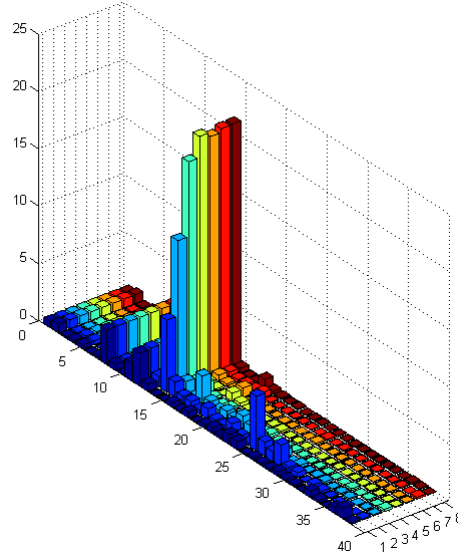


Figura 4.8: Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M, E_0)$.

Cálculo del $T\varphi(M_0, E)$

Los resultados del cálculo del espacio tangente $T\varphi(M_0, E)$ (etapa 3 del algoritmo) que caracteriza los inputs sensoriales que recibe un agente inmóvil ante el movimiento del entorno nos revelan que se necesitan 18 variables para describir el entorno (ver Fig. 4.9). Los correspondientes a 6 emisores de estímulos sensoriales en un espacio tridimensional.

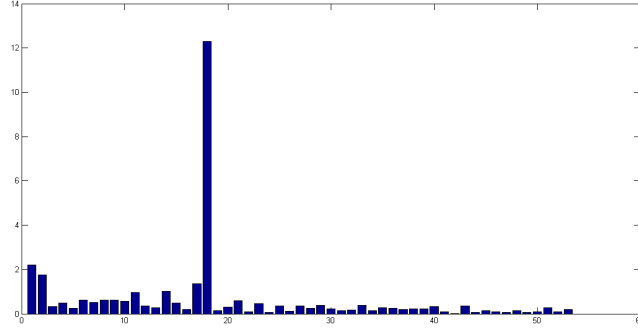


Figura 4.9: Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M_0, E)$.

Hay que darse cuenta de que algunos de los elementos del entorno también estimulaban de manera auditiva y táctil al agente siguiendo leyes psicofísicas diferentes. Sin embargo y pese a la diferente naturaleza de dichos estímulos, el algoritmo es capaz de distinguir claramente tanto los parámetros significativos del cuerpo (14 en total) como los emisores sensoriales (6 en total) que caracterizan el entorno.

Cálculo del $T\varphi(M_0, E) + T\varphi(M, E_0)$

Por último, obtenemos la dimensión del espacio suma $T\varphi(M_0, E) + T\varphi(M, E_0)$ que es el espacio tangente asociado a la variedad S cuando tanto el entorno como el cuerpo se mueven (etapa 4 del algoritmo). Como puede observarse (ver Fig. 4.10), el espacio suma es de dimensión 26.

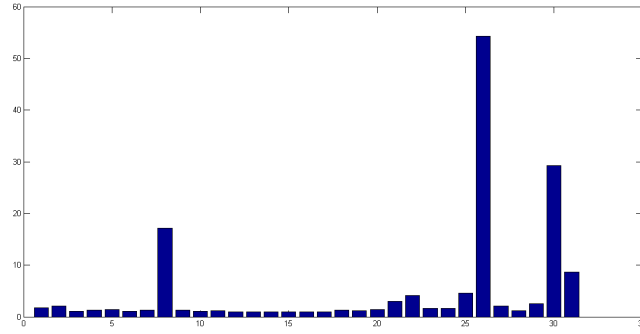


Figura 4.10: Resultados de la simulación correspondientes cálculo del espacio tangente $T\varphi(M_0, E) + T\varphi(M, E_0)$.

Cálculo de $T\varphi(M_0, E) \cap T\varphi(M, E_0)$

El espacio nulo de la concatenación de las bases de los espacios tangentes $T\varphi(M_0, E) + T\varphi(M, E_0)$ nos lleva a la obtención de una base del espacio intersección $T\varphi(M_0, E) \cap T\varphi(M, E_0)$ y de ahí a la extracción de una base generatriz del campo vectorial fundamental de S a partir de la base de $T\varphi(M_0, E) \cap T\varphi(M, E_0)$. La base del espacio intersección será de dimensión $14 + 18 - 26 = 6$ (ver etapa 4 del algoritmo).

Conclusiones del experimento

El resultado obtenido nos muestra que el agente percibe su espacio sensorimotor como un espacio de dimensión 6, es decir, que el número de posibles transformaciones compensables del sistema agente-entorno es 6. Este resultado es compatible con la configuración del experimento, porque sabemos que 6 es el número de transformaciones invariantes de un espacio de 3 dimensiones (matemáticamente esto es decir que el grupo de Lie de transformaciones ortogonales en un espacio de tres dimensiones es de dimensión 6). En este caso, estas 6 transformaciones invariantes que definen el espacio sensorimotor del agente, se corresponden a las 3 traslaciones y 3 rotaciones de un cuerpo en el espacio euclídeo de dimensión 3.

4.3. Diseño del entorno real y del agente físico

En esta sección, nos interesa demostrar la viabilidad del algoritmo de percepción sensorimotora en un entorno real. Las diferencias sustanciales con respecto al experimento anterior vendrán determinadas por la necesidad de trabajar con un agente más sencillo, con menor número de sensores y con restricciones en la dinámica 4.3 que gobierna los cambios del entorno. Además, por motivos de simplicidad, la captación de estímulos del entorno por parte del agente no se realizará en tiempo real (recogiendo la información sensorial a la vez que el agente se mueve) sino que éste realizará pequeños movimientos, parará, tomará medidas de sus sensores y proseguirá con el siguiente movimiento (de forma análoga a como se realiza en el experimento virtual). De esta forma se garantiza una captación de estímulos más precisa, al evitar errores externos e imprecisiones asociados al movimiento del propio agente.

El agente utilizado será un robot *Lego Mindstorms NXT* en un entorno acotado por el que se desplazará. Su movimiento vendrá generado por una serie de motores que le permitirá (i) desplazarse por el medio, y (ii) variar la orientación de sus dispositivos sensores. El agente podrá moverse libremente en su entorno físico teniendo acceso directo a los elementos emisores de estímulos sensoriales que éste le proporciona.

Se ha elegido utilizar la plataforma *Lego Mindstorms NXT* porque permite construir robots con configuraciones muy diferentes, de manera sencilla, disponemos de ellos e incluyen un conjunto bastante amplio de sensores de todo tipo: luminosos, sonoros, ultrasónicos, etc. La programación del robot se realizará a través del entorno *MATLAB*, ya que así se integra el control del mismo, tanto motor como sensor, junto con el núcleo matemático del algoritmo propuesto en una sola plataforma, simplificando el sistema. Para ello, se ha utilizado el paquete gratuito *RWTH - Mindstorms NXT Toolbox for MATLAB*, que permite un control en tiempo real del robot desde el entorno *MATLAB*.

4.3.1. Evaluación de la viabilidad experimental

En esta primera sección se muestran los resultados de las pruebas realizadas para evaluar si es viable el uso del sistema *Lego Mindstorms NXT*, además de un análisis de diferentes dispositivos sensoriales, con el fin de disponer de un agente físico integrado que cumpla condiciones técnicas para llevar a cabo los experimentos planteados.

A) EVALUACIÓN DEL SISTEMA MOTOR

Para evaluar la precisión de los motores *Lego Mindstorms NXT*, se construyó un sencillo vehículo con dos ruedas situadas en el mismo eje, cada una de ellas controladas por un motor, y se ejecutó una batería de pruebas consistente en realizar series de 10 giros de 180 grados. Como las condiciones del experimento que queremos plantear no requieren ningún tipo de restricción en el funcionamiento de los motores, para aumentar la precisión de éstos, se configuraron a una potencia del 20 % (suficiente para mover el robot), y se programó un modo de arranque suave (que evita que las ruedas patinen) y un modo de frenado que reduce suavemente la potencia del motor antes de completar el giro propuesto (aumentando considerablemente la precisión en el giro). Con esta batería de pruebas, se testeó que la precisión de giro de los motores con estos parámetros de configuración era de $\pm 1^\circ$ grado (ver Tabla. 4.1).

	EVALUACIÓN DEL SISTEMA MOTOR
Potencia motora	20 % de la potencia total del motor
Modo de arranque	SmoothStart en configuración motora RWTH
Modo de frenado	SmoothStart en configuración motora RWTH
Precisión	$\pm 1^\circ$ / giros 180 ^o

Tabla 4.1: Resumen de prestaciones y configuración del sistema motor.

Teniendo en cuenta que el algoritmo requiere pequeños movimientos del robot (que acumulan un error de giro mucho menor que movimientos amplios como de 180 grados), la precisión obtenida en las pruebas resultó satisfactoria y consideró el uso de la plataforma *Lego Mindstorms NXT* para la construcción de nuestro agente.

Implementación de prototipo

Construimos un primer prototipo de robot (ver Fig. 4.11) con unos parámetros de configuración motora que ofrecieran diversos grados de libertad, tanto para el desplazamiento del robot como para la orientación de los sensores, y que permitiera el uso de cualquier tipo de sensor (tanto los proporcionados por *Lego Mindstorms NXT* como otros disponibles) que se quisiera utilizar. Este prototipo, se diseña para que el agente físico pueda desplazarse mediante dos ruedas, cada una controlada por un motor, y dispone de dos brazos articulados en cuyos extremos se pueden colocar sensores cualesquiera. Los brazos articulados presentan dos movimientos de rotación, uno paralelo al suelo (izquierda-derecha) y otro perpendicular a él (arriba-abajo), cada uno controlado por un motor, que junto con las dos ruedas, dotan al robot de un número de grados de libertad motores lo suficientemente amplio como para poder realizar los experimentos planteados.

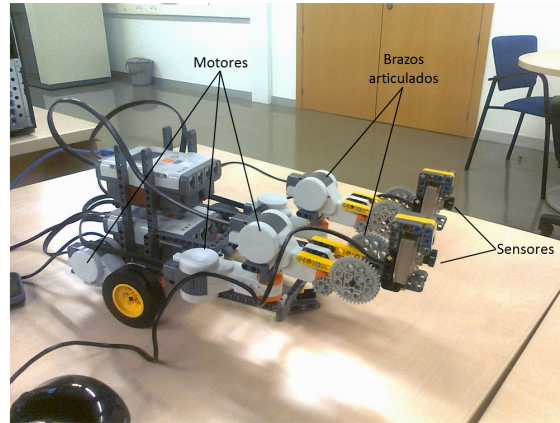


Figura 4.11: Primer prototipo de agente sensomotor. En este ejemplo en particular, los sensores colocados en los extremos de los brazos eran unas cámaras web.

B) EVALUACIÓN DE DISPOSITIVOS SENSORIALES

Una vez se dispone de un prototipo operativo del robot que actuará como agente, el siguiente paso es determinar el tipo de sensores que se utilizarán en el experimento, que a su vez condicionarán el tipo de entorno en el que interactuará el agente.

Dispositivos sensoriales *Lego*

Los sensores *Lego* para los cuales se puede construir un entorno compatible con el algoritmo a testar son el sensor ultrasónico y el sensor de sonido, que ofrecen las siguientes características:

SENSOR	RANGO	PRECISIÓN
Ultrasónico	0..255 cm	± 3 cm
Sonido	0..90 dB (en %)	1 %
Sistema motor (movimiento mínimo)	± 1 cm	~ 1 cm

Tabla 4.2: Resumen de precisiones de sensores y motores.

Las especificaciones de los sensores *Lego* utilizables no ofrecen un modo de funcionamiento lo suficientemente preciso para las necesidades del algoritmo, que requiere detectar pequeñas diferencias sensoriales debidas a los pequeños movimientos que ejecuta robot y, claramente (ver Tabla 4.2), ninguno de los sensores, ni el sonoro, ni el ultrasónico, se acerca al nivel de precisión requerido. El movimiento del prototipo con la precisión de giro motor estimada en 1° se traduce en desplazamientos mínimos de un cm. en la posición del sensor. Lo que descarta el sensor ultrasónico (precisión ± 3 cm.) y el sensor de sonido porque su representación en escala logarítmica (dB) no proporciona suficiente sensibilidad a pequeños movimientos de 1 cm.

Cámaras web tradicionales

Utilizar cámaras web como sensores permite recrear las condiciones del experimento virtual (al menos, a nivel visual. Ver sección 4.1.1) de manera directa, ya que puede modelar *la retina del agente* tal como se explicó en el experimento previo.

Cada cámara modela, en el agente físico, un “ojo” similar al del entorno virtual, utilizando una pequeña región de lo que captura la cámara y seleccionando un conjunto de píxeles al azar que actuarán como los sensores fotosensibles del entorno simulado (véase sección 4.1.1).

Este tipo de sensores exige un *entorno con unos emisores luminosos*. De esta forma, la configuración del experimento real se reduce a unas condiciones similares a las del experimento simulado.

El entorno ha de consistir en una serie de luces colocadas de manera arbitraria en el campo de visión de las cámaras web de manera que los píxeles seleccionados como receptores fotosensibles de la cámara detecten los valores de luminosidad provenientes de dichas luces (ver sección 4.1.1).



Figura 4.12: Ejemplo de prueba realizada con cámara web (a) y detalle del entorno (b).

Para testear la viabilidad de esta solución, se ejecutaron una serie de pruebas con una cámara web y una bombilla como única fuente luminosa del entorno para comprobar la precisión de los píxeles de la cámara como captadores de intensidad lumínica.

Las pruebas llevadas a cabo consistieron en situar la cámara a una distancia determinada de la bombilla (~ 30 cm) y variar ligeramente su posición, para comprobar la precisión del cambio de valor lumínico resultado de dichos movimientos (ver Fig. 4.12).

El resultado de estas pruebas resultó ser fallido. Los cambios medidos en la intensidad luminosa no se correspondían con los movimientos de la cámara. Incluso en condiciones estáticas las medidas registradas resultaban variables (resultado debido posiblemente a la iluminación variable en el tiempo de la bombilla). Se realizó el mismo tipo de prueba con diferentes modelos de cámaras web ofreciendo todas ellas un resultado similar (ver Tabla 4.3).

CÁMARAS	RESOLUCIÓN	PRECISIÓN
PlayStation Eye PS3	640×480 píxeles	—
Hercules Classic Silver	800×600 píxeles	—
Kinect Xbox 360	1280×960 píxeles	—

Tabla 4.3: Resumen de características de las cámaras probadas.

Las cámaras web no están diseñadas para obtener valores de intensidad luminosa. Por tanto, intentar obtener esta información a partir de las imágenes que devuelven, no resulta adecuado. Serían necesarios sensores lumínicos auténticos, específicamente diseñados para esta tarea. Ante la imposibilidad de disponer de dicho tipo de sensores, se plantearon nuevas alternativas.

Kinect: Sensor de profundidad

El sensor de profundidad de la Kinect proporciona imágenes de hasta 640×480 píxeles en donde cada uno de los píxeles contiene información en milímetros de la distancia del objeto que se proyecta sobre dicho pixel. El rango de funcionamiento del sensor de profundidad en el modo normal es de 80 cm a 5 m.

Para comprobar la precisión del sensor de profundidad de la Kinect, se realizaron una serie de pruebas (10 posiciones arbitrarias en torno a $1 \text{ m} \pm 20 \text{ cm}$, 10 posiciones en torno a $2 \text{ m} \pm 20 \text{ cm}$, y

10 posiciones en torno a $3 \text{ m} \pm 20 \text{ cm}$). Las pruebas consistieron en situar una caja enfrente de la cámara a diferentes distancias midiendo la diferencia entre la distancia real a la caja y la distancia medida por el sensor. El resumen de resultados se muestra en la Tabla 4.4.

DISTANCIA	ERROR MÁXIMO
$1 \text{ m} \pm 20 \text{ cm}$	2 mm
$2 \text{ m} \pm 20 \text{ cm}$	5 mm
$3 \text{ m} \pm 20 \text{ cm}$	2.5 cm

Tabla 4.4: Resumen de resultados de las pruebas de precisión con el sensor de profundidad de la Kinect.

Las pruebas con el sensor de profundidad de la Kinect resultaron satisfactorias para nuestros propósitos ya que la precisión del sensor de profundidad parece adecuada en términos de las condiciones de nuestro experimento (recordemos que el movimiento motor está en torno a 1 cm.) siempre que garanticemos que los elementos del entorno estén situados como máximo a 2 m de distancia del agente (ver Tabla 4.4).

Como resultado de esta primera fase de evaluación, tenemos garantizado que, nuestro prototipo cumple con las condiciones experimentales requeridas para poder reproducir el experimento virtual previo.

Además, plantea las restricciones en el diseño del entorno imprescindibles para que el agente físico disponga de información adecuada para la aplicación del algoritmo.

Diseño del entorno experimental

El entorno tendrá que estar formado por objetos que el robot pueda detectar utilizando su sensor de profundidad. Por ello, se decidió utilizar cajas de cartón, es decir, objetos que se pueden mover con facilidad y que son fáciles de detectar, sin errores, por su tamaño (ver Fig. 4.13).

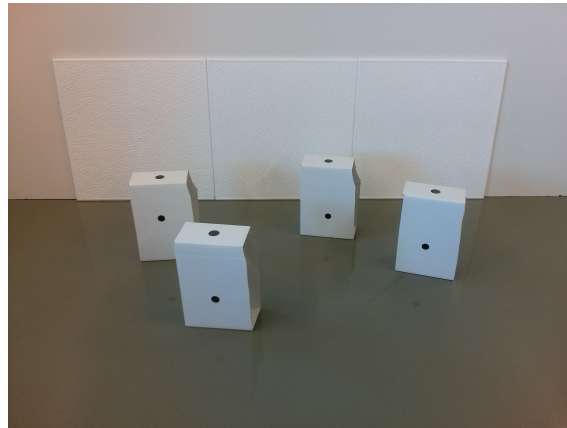


Figura 4.13: Detalle del entorno.

De esta forma, al igual que como se pretendía hacer con las cámaras web, ciertos píxeles de la Kinect harán las veces de sensores de la retina. En este caso, la información de intensidad lumínica del entorno con luces, será sustituida por la distancia de las cajas al robot, de una manera similar a como se hace en la simulación donde se toma la distancia de las luces a los ojos como medida de la intensidad luminosa que llega a la retina (ver sección 4.1.1).

En el experimento real, por tanto, las cajas distribuidas por el entorno hacen las veces de los emisores lumínicos del entorno virtual y la supuesta información luminosa vendrá dada por la distancia de las cajas al robot.

Esto no supone ningún problema puesto que el algoritmo funciona independientemente del tipo de estímulo y del modo en que éste llega a los dispositivos sensoriales. A efectos del algoritmo del agente, solo importan las relaciones sensomotoras agente-entorno, es decir, la información que puede extraer el agente sobre su entorno mediante su interacción sensomotora con él.

4.4. Experimento tipo 2: Entorno real

4.4.1. Configuración experimental

Las dimensiones y el peso del dispositivo sensorial seleccionado (Kinect) exigió cambios en el prototipo inicial del agente. Se decidió eliminar los brazos articulados del robot dejando un diseño mucho más simple en el que el único movimiento posible vendría dado por las ruedas (ver Fig. 4.14).



Figura 4.14: Robot con Kinect.

Sin embargo, esta configuración del agente presenta un problema importante dado que con el sensor de profundidad es imposible obtener (en cada pixel que tomamos como supuesto receptor de la retina) información de las distancias de todas las cajas, puesto que cada pixel solo devuelve la distancia al objeto que tiene delante.

El algoritmo requiere para funcionar que cada receptor de la retina registre el influjo de todos los elementos emisores de estímulos sensoriales del entorno. Sin embargo, es imposible con esta configuración del experimento recoger la información de todas las distancias a las cajas en cada uno de los píxeles que actúan como sensores (sólo se recoge información como máximo de un único elemento del entorno por pixel-receptor).

La solución de compromiso que se ha tomado ha sido la de obtener la información de las distancias a todas las cajas (que se corresponde con la intensidad lumínica en el ejemplo virtual) utilizando un algoritmo de visión por computador y simulando, posteriormente, su influjo sobre una retina simulada utilizando los valores reales obtenidos por el robot.

Aunque esta capa de software y virtualización extra elimina ciertas componentes reales del experimento, ha sido necesario tomar esta decisión de compromiso por cuestiones de aplicabilidad y de realización a partir de los recursos disponibles. Sin embargo, la esencia del experimento se mantiene, pues, aunque simplificado, se siguen obteniendo datos significativos extraídos de manera sensomotora como requiere el algoritmo.

4.4.2. Diseño experimental

A continuación se muestra en detalle el diseño del experimento real (ver Fig. 4.15), donde se indican (i) los parámetros utilizados en la configuración del entorno, (ii) la capa de virtualización extra simulando la proyección retiniana y la excitación sensorial en el dispositivo visual del agente físico, (iii) el sistema motor del robot y (iv) los sistemas nerviosos sensor y motor.

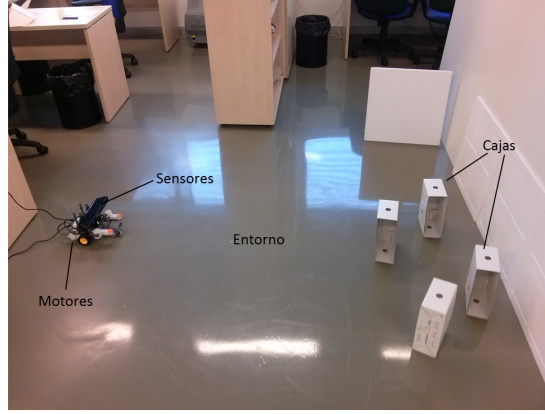


Figura 4.15: Entorno final de simulación.

i) Configuración del entorno

El entorno está constituido por cuatro cajas de cartón que realizan una función similar a las luces en la simulación como ya se ha explicado anteriormente. En el entorno simulado por ordenador, las luces estaban situadas a una distancia de entre 99 y 101 centímetros y separadas unas de otras entre 1 y 50 centímetros. El movimiento de las luces en la simulación es de como máximo 1 centímetro. Para el entorno real, situamos las cajas a una distancia de entre 150 y 190 centímetros de la posición inicial del robot y separadas unas de otras entre 20 y 50 centímetros. El movimiento de las cajas es de entre 5 y 20 centímetros. La precisión de la medición de las distancias a las cajas por parte de las cámaras se ha estimado en 5 mm, con lo que los movimientos de las cajas de mínimo 5 cm es para garantizar que cualquier cambio en la posición de las cajas sea detectado por el robot (ver Tabla 4.5).

	EXPERIMENTO VIRTUAL	EXPERIMENTO REAL
Distancia emisor-receptor	$2 \text{ m} \pm 20 \text{ cm}$	$2 \text{ m} \pm 20 \text{ cm}$
Distancia entre elementos	$[1..50] \text{ cm}$	$[20..50] \text{ cm}$
Precisión movimientos	$[0..1] \text{ cm}$	$[5..20] \text{ cm}$
Precisión estimación distancia	Absoluta (<i>MATLAB</i>)	5 mm

Tabla 4.5: Resumen de los parámetros de configuración del entorno en los experimentos virtual y real.

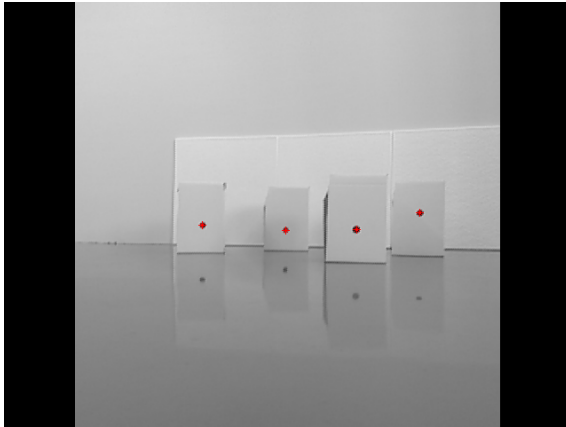
Adicionalmente, cada caja de cartón incluye un círculo negro para simplificar el algoritmo de detección de distancias, como se verá en la sección posterior.

ii) Configuración del dispositivo sensor del agente

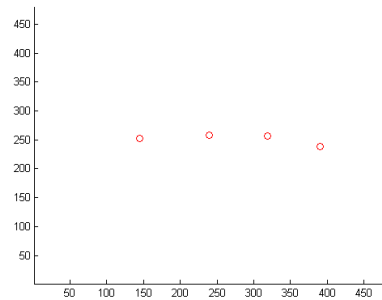
Para simular la proyección del entorno sobre la retina, se ha utilizado un algoritmo que (i) extrae las distancias del robot a las cajas que configuran el entorno, (ii) calcula sus proyecciones

teóricas sobre la retina y (iii) determina la excitación de cada sensor utilizando como simulación de intensidad luminosa la distancia a las cajas. Para calcular las distancias del robot a las cajas que configuran el entorno, se ha implementado un algoritmo de visión por computador (ver Fig. 4.16a) Tanto la cámara como el sensor de profundidad se han configurado para obtener imágenes de la misma resolución (640×480 píxeles) por ser la máxima que permite el sensor de profundidad. De esta manera, la correspondencia entre píxeles de cámara y sensor es directa.

1. Las fases del algoritmo de visión utilizado son las siguientes:
2. En primer lugar se *toma una foto con la cámara* (con el modo cámara de la Kinect) del entorno. Para facilitar la detección de las cajas, tanto éstas como el fondo del entorno son de color blanco con la particularidad de que cada una de las cajas tiene un círculo negro.
3. En segundo lugar *se umbraliza la imagen* y se aplica un algoritmo de conectividad para etiquetar cada círculo como un objeto diferente (que corresponde a una caja distinta en el entorno).
4. Se calculan las *coordenadas de sus centroides* (promedios de los píxeles fila y columna que forman el objeto).
5. A continuación, *se toma una imagen con el sensor de profundidad* de la Kinect.
6. *Para cada caja*, se utilizan las coordenadas de su centroide (obtenidas a partir de los círculos) como índice para buscar *su distancia en la imagen* de profundidades dada por la Kinect.



(a) Detalle del algoritmo de visión por computador. El algoritmo detecta las coordenadas del centroide de los círculos negros de las cajas (centroides marcados en rojo). Se puede observar en la imagen que el campo de visión de la cámara ha sido reducido de 640×480 a 480×480 píxeles para tener una correspondencia directa entre píxeles y coordenadas de la retina.



(b) Detalle de la proyección de los centroides sobre la retina en coordenadas de la cámara (tamaño 480×480) antes de hacer la transformación a coordenadas de la retina (tamaño 1×1).

Figura 4.16: Detalles de la simulación de la proyección de los elementos del entorno sobre la retina.

Ahora que ya se tienen las distancias de todas las cajas al centro de la cámara (que se utilizarán como medida de intensidad luminosa, falta simular la proyección de dichas intensidades (distancias) sobre los sensores de la retina.

La simulación de estas proyecciones se llevará a cabo de una manera similar a la del experimento virtual (ver sección 4.1.1). Para ello, estableceremos ciertas correspondencias (ver Tabla 4.6:

1. Supondremos una retina de tamaño 1×1 cm situada justo en el centro de la cámara al igual que en el experimento virtual.

2. La retina está constituida por 40 sensores situados aleatoriamente del mismo modo que en la simulación por ordenador.
3. Para calcular la posición sobre la cual se proyecta cada estímulo en la retina se representan las coordenadas en píxeles de cada caja (640×480 píxeles) en la escala de la retina (1×1 cm). La imagen de la cámara se recorta de 640×480 a 480×480 píxeles, para mantener una correspondencia con las proporciones en la retina (1×1 cm) (ver Fig 4.16b).
4. Para determinar la cantidad de luz que, a partir de esa proyección, le llega a cada sensor de la retina se utilizan las fórmulas de la ley psicofísica visual explicadas en la sección 4.1.1.

	EXPERIMENTO VIRTUAL	EXPERIMENTO REAL
Tamaño retina	1×1 cm	1×1 cm
Número de sensores por retina	40	40
Número total de sensores	90	40

Tabla 4.6: Resumen de los parámetros sensores de los experimentos virtual y real.

iii) Configuración del sistema motor del agente

El sistema motor de nuestro agente (ver sección 4.1.2) tiene cinco parámetros: Por cada rueda existe un parámetro que controla el ángulo de giro, y otro el sentido del mismo (4 parámetros, dos por rueda). Se añade un quinto parámetro que permite un desplazamiento adicional a una de las ruedas para ganar un movimiento extra en la rotación añadiendo algo de complejidad al movimiento del robot (ver Tabla 4.7).

El dispositivo sensor (cámara Kinect) se encuentra situado de manera rígida al cuerpo del robot sin añadir nuevos grados de libertad.

iv) Configuración del sistema nervioso sensor y motor del agente

El *sistema nervioso sensor* estará compuesto por 40 neuronas que reciben el estímulo de los 40 sensores totales del organismo (análogamente a la sección 4.1.3).

El *sistema nervioso motor* estará compuesto por 20 neuronas (número arbitrario para simular cierta complejidad neuronal motora) que se conectan con los elementos del “sistema efector-motor” (ver sección 4.1.2) dando lugar a una configuración concreta en los 5 parámetros del “sistema motor” del agente (véase Tabla 4.7).

	EXPERIMENTO VIRTUAL	EXPERIMENTO REAL
Número de neuronas sensoras	90	40
Número de neuronas motoras	40	20
Número de parámetros del sistema motor	16	5

Tabla 4.7: Resumen de las configuraciones motoras y nerviosas de los experimentos virtual y real.

4.4.3. Ejecución experimental y resultados

El esquema seguido es el mismo que el del experimento simulado (véase sección 4.2.1) pero teniendo en cuenta las siguientes consideraciones.

1. En la etapa en la que el robot realiza movimientos diferenciales (etapa 2), éste realiza una pausa de 0.5 segundos al final de cada movimiento diferencial antes de tomar los datos de

distancia. Esto se hace para evitar vibraciones de la cámara por efecto del movimiento del robot que puedan interferir en la calidad de las medidas. Además, al final de cada serie de movimientos diferenciales (es decir, tras cada movimiento completo, ver sección 4.2.1), el robot es recolocado manualmente de nuevo en la posición inicial para garantizar que los elementos del entorno se encuentren siempre en el campo visual de la cámara. Hay que recordar que la posición inicial de del robot en cada movimiento no tiene por qué ser la misma, pues lo significativo para el algoritmo son los cambios sensores mientras el robot se mueve.

2. Uno de los problemas es que el algoritmo necesita un elevado nivel de precisión para funcionar correctamente. Por esta razón redujimos el número de parámetros a aquellos para los que la precisión de las medidas estuviera garantizada. El robot ya quedó reducido a unos pocos movimientos por cuestiones del tamaño de la cámara. Utilizamos un número mínimo de cajas en el entorno para facilitar que todas las cajas estuvieran dentro del campo visual de la cámara. Las primeras pruebas utilizaban tres cajas en el entorno. Sin embargo, este escenario no ofrecía la suficiente complejidad para garantizar el funcionamiento del algoritmo. Se seleccionaron, por tanto, cuatro cajas como el mínimo número de elementos del entorno con el que realizar el experimento.

Resultados

Los resultados obtenidos en el experimento con el agente físico descrito fueron los siguientes:

Cálculo del $T\varphi(M, E_0)$

Como puede observarse en la Figura 4.17, a partir de una configuración de 20 neuronas controlando 5 parámetros motores, el número de parámetros que caracterizan la configuración del cuerpo del agente (sistema motor) queda reducida a 3 parámetros (precisamente los correspondientes a una rotación y dos translaciones), esto es, los tres grados de libertad en que puede moverse un vehículo de dos ruedas que se desplaza por el suelo como es nuestro robot.

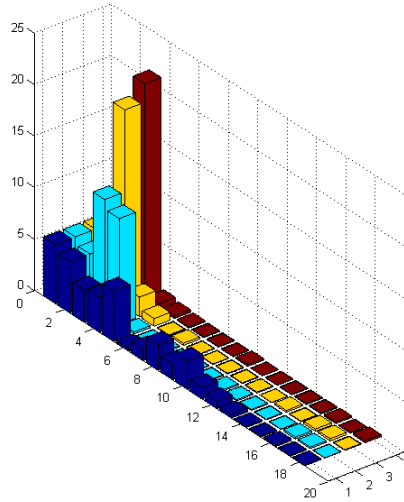
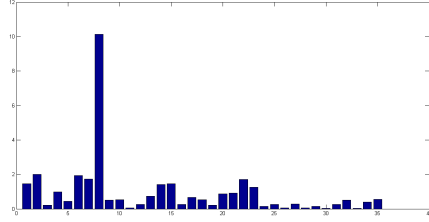


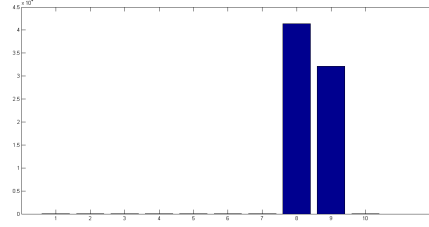
Figura 4.17: Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M, E_0)$. Como puede observarse en la figura, el número de iteraciones realizadas en esta etapa del algoritmo es de 4 iteraciones, frente a las 8 del experimento virtual (ver Anexo B). Esta reducción se programó así porque se vio en los resultados que 4 iteraciones eran necesarias para caracterizar el espacio $T\varphi(M, E_0)$ correctamente.

Cálculo del $T\varphi(M_0, E)$

Los resultados de la etapa 3 del algoritmo (ver Fig. 4.18a), que procesa la información de los estímulos sensoriales que recibe nuestro agente inmóvil ante el movimiento del entorno, revelan que se necesitan 8 variables para describir el entorno, las correspondientes a 4 cajas moviéndose en un espacio bidimensional.



(a) Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M_0, E)$.



(b) Resultados de la simulación correspondientes al cálculo del espacio tangente $T\varphi(M_0, E) + T\varphi(M, E_0)$.

Figura 4.18: Resultados de la simulación adicionales, correspondientes a las fases 3 y 4 del algoritmo.

Cálculo del $T\varphi(M_0, E) + T\varphi(M, E_0)$

Por último, con los resultados obtenidos, se calcula la dimensión de la suma de los dos subespacios obtenidos anteriormente (el espacio tangente asociado a la variedad S cuando tanto el entorno como el cuerpo se mueven). Como puede observarse en la Figura 4.18b, el espacio suma es de dimensión 8.

En este caso, la frontera en la determinación del número de valores singulares no nulos del espacio suma (ver Anexo B) no está tan marcada como en los resultados obtenidos en el experimento virtual. Esta pérdida de precisión se debe a la mayor complejidad experimental del entorno real, donde los resultados pueden verse afectados por multitud de factores externos (movimientos menos precisos, imprecisiones en el sensor, etc.) que no están presentes en el entorno virtual.

Cálculo del $T\varphi(M_0, E) \cap T\varphi(M, E_0)$

Siguiendo la fórmula de las dimensiones dada en la etapa 4 del algoritmo (ver sección 3.4), se obtiene que la intersección de los dos espacios tangentes estimados muestra un espacio sensomotor de dimensión $3 + 8 - 8 = 3$.

Conclusiones del experimento

El resultado obtenido nos muestra que el robot percibe su espacio sensomotor como un espacio de dimensión 3, es decir, que el número de posibles transformaciones compensables del sistema agente-entorno es 3. Este resultado es compatible con la configuración del experimento, pues estas 3 transformaciones invariantes que definen el espacio sensomotor del robot, se corresponden a las 2 traslaciones y 1 rotación en el espacio euclídeo de dimensión 2.

Capítulo 5

Conclusiones

5.1. Objetivos alcanzados

El objetivo de este proyecto era probar el modelo propuesto por [10, 11], según el cual un agente podría deducir las propiedades espaciales del mundo en que habita (lo que hemos denominado espacio sensomotor) a través de las relaciones que extrae entre sus movimientos y los estímulos sensoriales que recibe en respuesta.

Los resultados obtenidos tanto en el experimento realizado en el entorno virtual como en el entorno real muestran que existe un método viable que permite al agente obtener información sobre la estructura de su espacio sensomotor, concretamente las propiedades de simetría de las transformaciones invariantes (movimientos compensables) de su espacio sensomotor y su dimensión, a través de un conjunto de generadores motores.

5.2. Trabajo futuro

Ahora que disponemos de un método para obtener no solo la dimensión del espacio sensomotor de un agente, sino a las propiedades de invarianza y simetría de dicho espacio, una línea de trabajo futuro sería obtener, a partir de los datos obtenidos (el campo de vectores fundamental X^S asociado al conjunto $\Gamma^d(\varphi)$), información y propiedades adicionales acerca de la estructura del espacio sensomotor. Hasta el momento, hemos logrado obtener una base de generadores $\{V_m\}$ con la que generar una orden motora $M_U(t)$ y, con ello, el campo de vectores X^S que nos permite acceder a la estructura del espacio sensomotor. Sin embargo, esta base de generadores no es ortonormal. Una línea de trabajo cercana sería encontrar un modo de transformar esa base de generadores en una base ortonormal, lo cual nos permitiría calcular la conmutatividad de los movimientos compensables que definen el espacio sensomotor. En nuestro ejemplo experimental (ver sección 4.2.2), obteníamos un espacio sensomotor de dimensión 6, correspondiente a las 3 rotaciones y 3 traslaciones del espacio euclídeo de 3 dimensiones. Acceder a la conmutatividad del espacio sensomotor permitiría al agente moverse en términos de esas transformaciones (rotaciones y traslaciones) de manera independiente.

Una línea de trabajo más lejana sería intentar extraer otra clase de información del espacio sensomotor. De esta forma, se podría llegar a un método que permitiera al agente encontrar geodésicas del espacio sensomotor, que le permitieran moverse en “línea recta” (a través de su geodésica) en su entorno.

En cuanto a experimentos en el mundo físico, otra alternativa de trabajo sería testear el modelo propuesto por [11] en condiciones experimentales más complejas. Para ello se podrían utilizar robots con mayor número de grados de libertad actuando en entornos más complejos, con diferentes fuentes de estímulos diferentes naturalezas (lumínicas, sonoras, etc.) y con un agente con mayor número de sensores, más complejos y de diferentes tipos.

